

see listing A

08/24/99  
JCS03 U.S. PTO

PATENT  
Attorney Docket No. DIVER1180-1

JCS15 U.S. PTO  
09/382242  
08/24/99

☐ NEW PATENT APPLICATION  
☐ CONTINUATION-IN-PART  
☒ CONTINUATION  
☐ DIVISIONAL  
☐ FILE WRAPPER CONTINUATION

ASSISTANT COMMISSIONER  
FOR PATENTS  
Washington, D.C. 20231

|   |
|---|
| CERTIFICATE OF MAILING BY "EXPRESS MAIL"  |
| "EXPRESS MAIL" MAILING LABEL NO <u>EL382208665US</u>  |
| DATE OF DEPOSIT <u>August 24, 1999</u> HEREBY CERTIFY THAT THIS PAPER OR FEE IS BEING DEPOSITED WITH THE UNITED STATES POSTAL SERVICE "EXPRESS MAIL POST OFFICE TO ADDRESSEE" SERVICE UNDER 37 C.F.R. 1.10 ON THE DATE INDICATED ABOVE AND IS ADDRESSED TO THE ASSISTANT COMMISSIONER FOR PATENTS, WASHINGTON, D.C. 20231 |
| <u>Rita H. Jennings</u><br>(TYPED OR PRINTED NAME OF PERSON MAILING PAPER OR FEE)   |
| <u>Rita H. Jennings</u><br>(SIGNATURE OF PERSON MAILING PAPER OR FEE)   |

Sir:

Transmitted herewith for filing is the divisional patent application of

Inventors: Dan E. Robertson; Dennis Murphy; John Reid; Anthony M. Maffia; Steven Link;  
Ronald V. Swanson; Patrick V. Warren

For: **ESTERASES**

This is a request for filing a X continuation ☐ divisional application under 37 C.F.R. 1.53(b), of prior Application No. 08/602,359, filed on February 16, 1996, now pending.

|                              |  |   |                             |
|------------------------------|--|---|-----------------------------|
| FULL NAME OF FIRST INVENTOR  | LAST NAME:<br><b>Robertson</b>                   | FIRST NAME:<br><b>Dan</b>                         | MIDDLE NAME:<br><b>E.</b>   |
| CITIZENSHIP                  | STATE OR FOREIGN COUNTRY: <b>USA</b>             |   |                             |
| POST OFFICE ADDRESS          | POST OFFICE ADDRESS:<br><b>33 Evergreen Lane</b> | CITY AND STATE:<br><b>Haddonfield, New Jersey</b> | ZIP CODE:<br><b>08033</b>   |
| FULL NAME OF SECOND INVENTOR | LAST NAME:<br><b>Murphy</b>                      | FIRST NAME:<br><b>Dennis</b>                      | MIDDLE NAME:<br><b>none</b> |
| CITIZENSHIP                  | STATE OR FOREIGN COUNTRY: <b>USA</b>             |   |                             |
| POST OFFICE ADDRESS          | POST OFFICE ADDRESS:<br><b>10 Fairway Road</b>   | CITY AND STATE:<br><b>Paoli, Pennsylvania</b>     | ZIP CODE:<br><b>19301</b>   |

|                               |  |  |                             |
|-------------------------------|--|--|-----------------------------|
| FULL NAME OF THIRD INVENTOR   | LAST NAME:<br><b>Reid</b>                                      | FIRST NAME:<br><b>John</b>                           | MIDDLE NAME:<br><b>none</b> |
| CITIZENSHIP                   | STATE OR FOREIGN COUNTRY: <b>USA</b>                           |  |                             |
| POST OFFICE ADDRESS           | POST OFFICE ADDRESS:<br><b>922 Montgomery Avenue, Apt. J-2</b> | CITY AND STATE:<br><b>Bryn Mawr, Pennsylvania</b>    | ZIP CODE:<br><b>19010</b>   |
| FULL NAME OF FOURTH INVENTOR  | LAST NAME:<br><b>Maffia</b>                                    | FIRST NAME:<br><b>Anthony</b>                        | MIDDLE NAME:<br><b>M.</b>   |
| CITIZENSHIP                   | STATE OR FOREIGN COUNTRY: <b>USA</b>                           |  |                             |
| POST OFFICE ADDRESS           | POST OFFICE ADDRESS:<br><b>2505-3A Cedar Tree Lane</b>         | CITY AND STATE:<br><b>Wilmington, Delaware</b>       | ZIP CODE:<br><b>19810</b>   |
| FULL NAME OF FIFTH INVENTOR   | LAST NAME:<br><b>Link</b>                                      | FIRST NAME:<br><b>Steven</b>                         | MIDDLE NAME:<br><b>none</b> |
| CITIZENSHIP                   | STATE OR FOREIGN COUNTRY: <b>USA</b>                           |  |                             |
| POST OFFICE ADDRESS           | POST OFFICE ADDRESS:<br><b>108 Gibson Avenue</b>               | CITY AND STATE:<br><b>Wilmington, Delaware</b>       | ZIP CODE:<br><b>19803</b>   |
| FULL NAME OF SIXTH INVENTOR   | LAST NAME:<br><b>Swanson</b>                                   | FIRST NAME:<br><b>Ronald</b>                         | MIDDLE NAME:<br><b>V.</b>   |
| CITIZENSHIP                   | STATE OR FOREIGN COUNTRY: <b>USA</b>                           |  |                             |
| POST OFFICE ADDRESS           | POST OFFICE ADDRESS:<br><b>309 N. Lemon Street, Apt. A</b>     | CITY AND STATE:<br><b>Media, Pennsylvania</b>        | ZIP CODE:<br><b>19063</b>   |
| FULL NAME OF SEVENTH INVENTOR | LAST NAME:<br><b>Warren</b>                                    | FIRST NAME:<br><b>Patrick</b>                        | MIDDLE NAME:<br><b>V.</b>   |
| CITIZENSHIP                   | STATE OR FOREIGN COUNTRY: <b>USA</b>                           |  |                             |
| POST OFFICE ADDRESS           | POST OFFICE ADDRESS:<br><b>Sheffield Ave.</b>                  | CITY AND STATE:<br><b>Philadelphia, Pennsylvania</b> | ZIP CODE:<br><b>19163</b>   |
| FULL NAME OF EIGHTH INVENTOR  | LAST NAME:<br><b>Kosmotka</b>                                  | FIRST NAME:<br><b>Anna</b>                           | MIDDLE NAME:<br><b>none</b> |
| CITIZENSHIP                   | STATE OR FOREIGN COUNTRY: <b>USA</b>                           |  |                             |
| POST OFFICE ADDRESS           | POST OFFICE ADDRESS:<br><b>224 Harrison Road</b>               | CITY AND STATE:<br><b>Brookhaven, Pennsylvania</b>   | ZIP CODE:<br><b>19015</b>   |

No payment of the issue fee, abandonment of, or termination of proceeding has occurred in the above-identified prior application.

1. ☒ Cancel in this application original claims 2-20 of the prior application. (At least one original independent claim must be retained for filing purposes.)
2. ☒ A preliminary amendment is enclosed.

The filing fee has been calculated as shown below:

| For   | Number<br>Filed |   | Number<br>Extra |   | Rate         |              |   | Fee             |                 |
|---|-----------------|---|-----------------|---|--------------|--------------|---|-----------------|-----------------|
|   |                 |   |                 |   | Small Entity | Other Entity |   | Small<br>Entity | Other<br>Entity |
| Total Claims  | 6               | = | 0               | X | \$9          | \$18         | = | \$ .00          | \$ 0            |
| Independent Claims  | 1               | = | 0               | X | \$39         | \$78         | = | \$ .00          | 0               |
| Multiple Dependent<br>Claims Presented: <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No |                 |   |                 |   | \$130        | \$260        |   |                 | 0               |
| BASIC FEE   |                 |   |                 |   | \$380        | \$760        |   | \$380.00        | \$ 0            |
|   |                 |   |                 |   | TOTAL FEE    |              |   | \$380.00        | \$ 0            |

3. ☒ Please charge my Deposit Account No. 07-1895 the TOTAL FEE of \$380.00, which covers the filing fee for this application. A duplicate copy of this sheet is enclosed.
4. ☒ The Assistant Commissioner is hereby authorized to charge payment of the following fees associated with this communication or credit any overpayment to Deposit Account No. 07-1895. A duplicate copy of this sheet is enclosed.
  - ☒ Any additional filing fees required under 37 C.F.R. 1.16.
  - ☒ Any patent application processing fees under 37 C.F.R. 1.17.
5. ☒ Amend the specification by inserting before the first paragraph on page 1:  
  
This application is a ☒ continuation ☐ divisional of application Serial No. 08/602,359 filed on February 16, 1996, now pending; the entire contents of which are hereby incorporated by reference herein.
6. ☒ A verified statement claiming small entity status was filed in parent application Serial No. 08/602,359, filed July 25, 1996, and such status is still proper.
7. ☒ The prior application is assigned of record to RECOMBINANT BIOCATALYSIS, INC.
8. ☒ The power of attorney in the prior application is to Lisa A. Haile, Registration No. 38,347.
9. ☒ Please transfer the drawings from the prior application to the new application.

In re Application of  
Robertson et al.  
Page 4

Attorney Docket No.: DIVER1180-1

10. X A true copy of the prior application as filed is enclosed, including the Declaration and Power of Attorney filed in parent application, U.S. Serial No. 08/602,359, filed February 16, 1996.
11. X An Associate Power of Attorney is enclosed.
12. — Information Disclosure Statements filed in the prior application under 37 C.F.R. 1.97 are hereby made of record.
13. X Please transfer the computer readable form (CRF) copy of the Sequence Listing from the prior application, which CRF copy was filed with a Communication mailed July 28, 1997, to this new application.
14. X Please transfer the Statement under 37 C.F.R. § 1.821(f) and (g) from the prior application, which Statement was filed with a Communication mailed July 28, 1997, to this new application.
15. — Also enclosed: Copy of Petition for Extension of Time in parent application U.S. Serial No.: \_\_\_\_\_

Address all future communications to:

Lisa A. Haile, Ph.D.  
GRAY CARY WARE & FREIDENRICH LLP  
4365 Executive Drive, Suite 1600  
San Diego, California 92121-2189  
Telephone: 858-677-1456  
Facsimile: 858-677-1465

The undersigned states that the enclosed application papers comprise a true copy of the prior application as filed.

Respectfully submitted,

Date: August 24, 1999



Lisa A. Haile, Ph.D.  
Attorney for Applicant  
Registration No. 38,347

GRAY CARY WARE & FREIDENRICH LLP  
4365 Executive Drive, Suite 1600  
San Diego, CA 92121-2189  
GT6143399  
104703-156621

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: ) Group Art Unit: (Unassigned)  
Robertson et al. )  
 ) Examiner: (Unassigned)  
Filed: Herewith )  
 )  
Parent Serial No.: 08/602,359 )  
 )  
Parent Filing Date: February 16, 1996 )  
 )  
For: ESTERASES )  
 )  
 )  
 )

Box Patent Application  
Assistant Commissioner for Patents  
Washington, D.C. 20231

**PRELIMINARY AMENDMENT**

Sir:

This Preliminary Amendment is being filed herewith further to a request under 37 C.F.R. § 1.53(b) to file a continuation application based on Application Serial No. 08/602,359, filed February 16, 1996, now pending.

Please cancel claim 1 of the application, and add new claims 21-26 as follows:

--21. (New) An oligonucleotide probe consisting of at least about 15 contiguous nucleotides of a polynucleotide selected from the group consisting of SEQ ID NO:23-31 and SEQ ID NO:32.

22. (New) An oligonucleotide probe fully complementary to an oligonucleotide probe of Claim 21.

In re Application of:

Robertson et al.

Application No.: Unassigned

Filed: Herewith

Page 2

23. (New) The oligonucleotide probe of claims 21 or 22 wherein the probe is 20-50 nucleotides in length.

24. (New) The oligonucleotide probe of claims 21 or 22 wherein the probe is labeled with a detectable label.

25. (New) The oligonucleotide probe of claim 24, wherein the detectable label is an isotopic label or a non-isotopic label, which non-isotopic label is selected from the group consisting of: a fluorescent molecule, a chemiluminescent molecule, an enzyme, a cofactor, an enzyme substrate, and a hapten.

26. (New) The oligonucleotide probe of Claim 24, wherein the probe comprises a sequence which specifically hybridizes to a nucleic acid comprising SEQ ID NO:23-32 or a sequence fully complementary thereto to form a detectable target probe duplex.--

### Remarks

By the present communication, new claims 21-26 have been added. No new matter is introduced by the new claim language, as the newly presented claims are fully supported by Applicant's specification and original claims. Accordingly, claims 21-26 are currently pending.

It is believed that the application is in condition for allowance and, therefore, prompt and favorable action is earnestly solicited. If there are any questions concerning this communication, the Examiner is invited to call the undersigned at the telephone number provided below.

In re Application of:

Robertson et al.

Application No.: Unassigned

Filed: Herewith

Page 3

No fee is deemed necessary in connection with the filing of this Preliminary Amendment. However, if any fee is required, authorization is given to charge the amount of this fee to Deposit Account No. 07-1895.

Respectfully submitted,

Date: August 24, 1999

Lisa A. Haile, Ph.D.

Telephone: (858) 677-1456

Facsimile: (858) 677-1465

GRAY, CARY, WARE & FREIDENRICH, LLP  
4365 Executive Drive, Suite 1600  
San Diego, California 92121-2189

## ESTERASES

This invention relates to newly identified polynucleotides, polypeptides encoded by such polynucleotides, the use of such polynucleotides and polypeptides, as well as the production and isolation of such polynucleotides and polypeptides. More particularly, the polynucleotides and polypeptides of the present invention have been putatively identified as esterases. Esterases are enzymes that catalyze the hydrolysis of ester groups to organic acids and alcohols.

Many esterases are known and have been discovered in a broad variety of organisms, including bacteria, yeast and higher animals and plants. A principal example of esterases are the lipases, which are used in the hydrolysis of lipids, acidolysis(replacement of an esterified fatty acid with a free fatty acid) reactions, transesterification(exchange of fatty acids between triglycerides)reactions, and in ester synthesis. The major industrial applications for lipases include: the detergent industry, where they are employed to decompose fatty materials in laundry stains into easily removable hydrophilic substances; the food and beverage industry where they are used in the manufacture of cheese, the ripening and flavoring of cheese, as antistaling agents for bakery products, and in the production of margarine and other spreads with natural



butter flavors; in waste systems; and in the pharmaceutical industry where they are used as digestive aids.

The polynucleotides and polypeptides of the present invention have been identified as esterases as a result of their enzymatic activity.

In accordance with one aspect of the present invention, there are provided novel enzymes, as well as active fragments, analogs and derivatives thereof.

In accordance with another aspect of the present invention, there are provided isolated nucleic acid molecules encoding the enzymes of the present invention including mRNAs, cDNAs, genomic DNAs as well as active analogs and fragments of such enzymes.

In accordance with another aspect of the present invention there are provided isolated nucleic acid molecules encoding mature polypeptides expressed by the DNA contained in ATCC Deposit No. \_\_\_\_\_.

In accordance with yet a further aspect of the present invention, there is provided a process for producing such polypeptides by recombinant techniques comprising culturing recombinant prokaryotic and/or eukaryotic host cells, containing a nucleic acid sequence of the present invention, under conditions promoting expression of said enzymes and subsequent recovery of said enzymes.

In accordance with yet a further aspect of the present invention, there is provided a process for utilizing such enzymes, or polynucleotides encoding such enzymes for hydrolyzing ester groups to yield an organic acid and an alcohol. The esterases of the invention are stable at high temperatures and in organic solvents and, thus, are superior for use in production of optically pure chiral compounds used in pharmaceutical, agricultural and other chemical industries.

In accordance with yet a further aspect of the present invention, there are also provided nucleic acid probes comprising nucleic acid molecules of sufficient length to hybridize to a nucleic acid sequence of the present invention.

In accordance with yet a further aspect of the present invention, there is provided a process for utilizing such enzymes, or polynucleotides encoding such enzymes, for *in vitro* purposes related to scientific research, for example, to generate probes for identifying similar sequences which might encode similar enzymes from other organisms by using certain regions, *i.e.*, conserved sequence regions, of the nucleotide sequence.

These and other aspects of the present invention should be apparent to those skilled in the art from the teachings herein.

The following drawings are illustrative of embodiments of the invention and are not meant to limit the scope of the invention as encompassed by the claims.

Figure 1 is an illustration of the full-length DNA (SEQ ID NO:23) and corresponding deduced amino acid sequence (SEQ ID NO:33) of *Staphylothermus marinus* F1-12LC of the present invention. Sequencing was performed using a 378 automated DNA sequencer (Applied Biosystems, Inc.) for all sequences of the present invention.

Figure 2 is an illustration of the full-length DNA (SEQ ID NO:24) and corresponding deduced amino acid sequence (SEQ ID NO:34) of *Pyrodictium* TAG11-17LC.

Figure 3 is an illustration of the full-length DNA (SEQ ID NO:25) and corresponding deduced amino acid sequence (SEQ ID NO:35) of *Archaeoglobus venificus* SNP6-24LC.

Figure 4 is an illustration of the full-length DNA (SEQ ID NO:26) and corresponding deduced amino acid sequence (SEQ ID NO:36) of *Aquifex pyrophilus*-28LC.

Figure 5 is an illustration of the full-length DNA (SEQ ID NO:27) and corresponding deduced amino acid sequence (SEQ ID NO:37) of M11TL-29L.

Figure 6 is an illustration of the full-length DNA (SEQ ID NO:28) and corresponding deduced amino acid sequence (SEQ ID NO:38) of *Thermococcus* CL-2-30LC.

Figure 7 is an illustration of the full-length DNA (SEQ ID NO:29) and corresponding deduced amino acid sequence (SEQ ID NO:39) of *Aquifex* VF5-34LC.

Figure 8 is an illustration of the full-length DNA (SEQ ID NO:30) and corresponding deduced amino acid sequence (SEQ ID NO:40) of *Teredinibacter*-42L.

Figure 9 is an illustration of the full-length DNA (SEQ ID NO:31) and corresponding deduced amino acid sequence (SEQ ID NO:41) of *Archaeoglobus fulgidus* VC16-16MC.

Figure 10 is an illustration of the full-length DNA (SEQ ID NO:32) and corresponding deduced amino acid sequence (SEQ ID NO:42) of *Sulfolobus solfataricus* P1-8LC.

The term "gene" means the segment of DNA involved in producing a polypeptide chain; it includes regions preceding and following the coding region (leader and trailer) as well as intervening sequences (introns) between individual coding segments (exons).

A coding sequence is "operably linked to" another coding sequence when RNA polymerase will transcribe the two coding sequences into a single mRNA, which is then translated into a single polypeptide having amino acids derived from both coding sequences. The coding sequences need not be contiguous to one another so long as the expressed sequences ultimately process to produce the desired protein.

"Recombinant" enzymes refer to enzymes produced by recombinant DNA techniques; *i.e.*, produced from cells transformed by an exogenous DNA construct encoding the desired enzyme. "Synthetic" enzymes are those prepared by chemical synthesis.

A DNA "coding sequence of" or a "nucleotide sequence encoding" a particular enzyme, is a DNA sequence which is transcribed and translated into an enzyme when placed under the control of appropriate regulatory sequences.

In accordance with an aspect of the present invention, there are provided isolated nucleic acids (polynucleotides) which encode for the mature enzymes having the deduced amino acid sequences of Figures 1-10 (SEQ ID NOS:23-32).

In accordance with another aspect of the present invention, there are provided isolated polynucleotides encoding the enzymes of the present invention. The deposited material is a mixture of genomic clones comprising DNA encoding an enzyme of the present invention. Each genomic clone comprising the respective DNA has been inserted into a pBluescript vector (Stratagene, La Jolla, CA). The deposit has been deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland 20852, USA, on December 13, 1995 and assigned ATCC Deposit No. \_\_\_\_\_.

The deposit(s) have been made under the terms of the Budapest Treaty on the

International Recognition of the deposit of micro-organisms for purposes of patent procedure. The strains will be irrevocably and without restriction or condition released to the public upon the issuance of a patent. These deposits are provided merely as convenience to those of skill in the art and are not an admission that a deposit would be required under 35 U.S.C. §112. The sequences of the polynucleotides contained in the deposited materials, as well as the amino acid sequences of the polypeptides encoded thereby, are controlling in the event of any conflict with any description of sequences herein. A license may be required to make, use or sell the deposited materials, and no such license is hereby granted.

The polynucleotides of this invention were originally recovered from genomic gene libraries derived from the following organisms:

*Staphylothermus marinus* F1 is a thermophilic sulfur archaea which was isolated in Vulcano, Italy. It grows optimally at 85°C ( $T_{\max} = 98^{\circ}\text{C}$ ) at pH 6.5.

*Pyrodictium* TAG11 is a thermophilic sulfur archaea which was isolated in the Middle Atlantic Ridge. It grows optimally at 103°C ( $T_{\max} = 110^{\circ}\text{C}$ ) at pH 6.5.

*Archaeoglobus venificus* SNP6 was isolated in the Middle Atlantic Ridge and grows optimally at 75°C ( $T_{\max} = 92^{\circ}\text{C}$ ) at pH 6.9.

*Aquifex pyrophilus* K01 5a was isolated at Kolbeinsey Ridge, North of Iceland. This marine organism is a gram-negative, rod-shaped, strictly chemolithoautrophic, knall gas bacterium. It grows optimally at 85°C ( $T_{\max} = 95^{\circ}\text{C}$ ) at pH 6.8.

M11TL is a new species of *Desulfurococcus* which was isolated from Diamond Pool (formerly Jim's Black Pool) in Yellowstone. The organism grows heterotrophically by fermentation of different organic materials (sulfur is not necessary)

in grape-like aggregates optimally at 85 - 88°C in a low salt medium at pH 7.0.

*Thermococcus* CL-2 was isolated in the North Cleft Segment of the Juan de Fuca Ridge from a severed alvinellid worm residing on a "black smoker" sulfide structure. This marine archaea forms pleomorphic cocci, and grows optimally at 88°C.

*Aquifex* VF5 was isolated at a beach in Vulcano, Italy. This marine organism is a gram-negative, rod-shaped, strictly chemolithoautotrophic, knall gas bacterium. It grows optimally at 85°C ( $T_{\max} = 95^{\circ}\text{C}$ ) at pH 6.8.

*Teredinibacter* (pure) is an endosymbiont of the shipworm *Bankia gouldi*. The organism has straight to slightly bent 5-10  $\mu\text{m}$  rods, and forms spiral cells as stationary phase is met. The organism was described in Science (1983) 22:1401-1403. It grows optimally at 30°C at pH 8.0.

*Archaeoglobus fulgidus* VC16 was isolated in Vulcano, Italy. The organism grows optimally at 85°C ( $T_{\max} = 92^{\circ}\text{C}$ ) at pH 7.0.

*Sulfolobus solfataricus* P1 grows optimally at 85°C ( $T_{\max} = 87^{\circ}\text{C}$ ) at pH 2.0.

Accordingly, the polynucleotides and enzymes encoded thereby are identified by the organism from which they were isolated, and are sometimes hereinafter referred to as F1/12LC (Figure 1 and SEQ ID NOS:23 and 33), TAG11/17LC (Figure 2 and SEQ ID NOS:24 and 34), SNP6/24LC (Figure 3 and SEQ ID NOS:25 and 35), AqP/28LC (Figure 4 and SEQ ID NOS:26 and 36), M11TL/29L (Figure 5 and SEQ ID NOS:27 and 37), CL-2/30LC (Figure 6 and SEQ ID NOS:28 and 38), VF5/34LC (Figure 7 and SEQ ID NOS:29 and 39), Trb/42L (Figure 8 and SEQ ID NOS:30 and 40), VC16/16MC (Figure 9 and SEQ ID NOS:31 and 41) and P1/8LC (Figure 10 and SEQ ID NOS: 32 and 42).

The polynucleotides and polypeptides of the present invention show identity at the nucleotide and protein level to known genes and proteins encoded thereby as shown in Table 1.

Table 1

| Enzyme     | Gene w/closest Homology (Organism)   | Protein Similarity (%) | Protein Identity (%) | DNA Identity (%) |
|------------|--|------------------------|----------------------|------------------|
| F1/12LC    | No significant homology  | -                      | -                    | -                |
| TAG11/17LC | No significant homology  | -                      | -                    | -                |
| SNP6/24LC  | PIR S34609 - carboxylesterase <i>Pseudomonas</i> sp. (strain KWI-56) open reading frame of unknown function in <i>E.coli</i> . | 46                     | 27                   | 42               |
| AqP/29LC   |  | 53                     | 31                   | 38               |
| M11TL/29LC | No significant homology  | -                      | -                    | -                |
| CL02/30LC  | No significant homology  | -                      | -                    | -                |
| VF5/34LC   | Identified by homology to 28LC; also homologous to ORF of unknown function 5' of tgs in <i>E. coli</i>                         | 84                     | 71                   | 71               |
| Trb/42L    | No significant homology  | -                      | -                    | -                |
| P1-8LC     |  |                        |                      |                  |
| VC16-16MC  |  |                        |                      |                  |

All the clones identified in Table 1 encode polypeptides which have esterase activity.

This invention, in addition to the isolated nucleic acid molecules encoding the enzymes of the present invention, also provides substantially similar sequences. Isolated nucleic acid sequences are substantially similar if: (i) they are capable of hybridizing under conditions hereinafter described, to the polynucleotides of SEQ ID NOS:23-32; (ii) or they encode DNA sequences which are degenerate to the polynucleotides of SEQ ID NOS:23-32. Degenerate DNA sequences encode the amino acid sequences of SEQ ID NOS:33-42, but have variations in the nucleotide coding sequences. As used herein, substantially similar refers to the sequences having similar identity to the sequences of the instant invention. The nucleotide sequences that are substantially the same can be identified by hybridization or by sequence comparison. Enzyme sequences that are substantially the same can be identified by one or more of the following: proteolytic digestion, gel electrophoresis and/or microsequencing.

One means for isolating the nucleic acid molecules encoding the enzymes of the present invention is to probe a gene library with a natural or artificially designed probe using art recognized procedures (see, for example: Current Protocols in Molecular Biology, Ausubel F.M. *et al.* (EDS.) Green Publishing Company Assoc. and John Wiley Interscience, New York, 1989, 1992). It is appreciated by one skilled in the art that the polynucleotides of SEQ ID NOS:23-32, or fragments thereof (comprising at least 12 contiguous nucleotides), are particularly useful probes. Other particularly useful probes for this purpose are hybridizable fragments of the sequences of SEQ ID NOS:1-22 (*i.e.*, comprising at least 12 contiguous nucleotides).

With respect to nucleic acid sequences which hybridize to specific nucleic acid sequences disclosed herein, hybridization may be carried out under conditions of reduced stringency, medium stringency or even stringent conditions. As an example of oligonucleotide hybridization, a polymer membrane containing immobilized denatured nucleic acids is first prehybridized for 30 minutes at 45°C in a solution consisting of 0.9 M NaCl, 50 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 7.0, 5.0 mM Na<sub>2</sub>EDTA, 0.5% SDS, 10X Denhardt's, and 0.5 mg/mL polyriboadenylic acid. Approximately 2 X 10<sup>7</sup> cpm (specific activity 4-



9 X 10<sup>8</sup> cpm/ug) of <sup>32</sup>P end-labeled oligonucleotide probe are then added to the solution.

After 12-16 hours of incubation, the membrane is washed for 30 minutes at room temperature in 1X SET (150 mM NaCl, 20 mM Tris hydrochloride, pH 7.8, 1 mM Na<sub>2</sub>EDTA) containing 0.5% SDS, followed by a 30 minute wash in fresh 1X SET at T<sub>m</sub> -10°C for the oligo-nucleotide probe. The membrane is then exposed to autoradiographic film for detection of hybridization signals.

Stringent conditions means hybridization will occur only if there is at least 90% identity, preferably at least 95% identity and most preferably at least 97% identity between the sequences. See J. Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual, 2d Ed.*, Cold Spring Harbor Laboratory (1989) which is hereby incorporated by reference in its entirety.

As used herein, a first DNA (RNA) sequence is at least 70% and preferably at least 80% identical to another DNA (RNA) sequence if there is at least 70% and preferably at least a 80% or 90% identity, respectively, between the bases of the first sequence and the bases of the another sequence, when properly aligned with each other, for example when aligned by BLASTN.

The present invention relates to polynucleotides which differ from the reference polynucleotide such that the changes are silent changes, for example the change do not alter the amino acid sequence encoded by the polynucleotide. The present invention also relates to nucleotide changes which result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference polynucleotide. In a preferred aspect of the invention these polypeptides retain the same biological action as the polypeptide encoded by the reference polynucleotide.

The polynucleotides of this invention were recovered from genomic gene libraries from the organisms listed in Table 1. Gene libraries were generated in the Lambda ZAP II cloning vector (Stratagene Cloning Systems). Mass excisions were

performed on these libraries to generate libraries in the pBluescript phagemid. Libraries were generated and excisions were performed according to the protocols/methods hereinafter described.

The polynucleotides of the present invention may be in the form of RNA or DNA which DNA includes cDNA, genomic DNA, and synthetic DNA. The DNA may be double-stranded or single-stranded, and if single stranded may be the coding strand or non-coding (anti-sense) strand. The coding sequences which encodes the mature enzymes may be identical to the coding sequences shown in Figures 1-10 (SEQ ID NOS:23-32) or may be a different coding sequence which coding sequence, as a result of the redundancy or degeneracy of the genetic code, encodes the same mature enzymes as the DNA of Figures 1-10 (SEQ ID NOS:23-32).

The polynucleotide which encodes for the mature enzyme of Figures 1-10 (SEQ ID NOS:33-42) may include, but is not limited to: only the coding sequence for the mature enzyme; the coding sequence for the mature enzyme and additional coding sequence such as a leader sequence or a proprotein sequence; the coding sequence for the mature enzyme (and optionally additional coding sequence) and non-coding sequence, such as introns or non-coding sequence 5' and/or 3' of the coding sequence for the mature enzyme.

Thus, the term "polynucleotide encoding an enzyme (protein)" encompasses a polynucleotide which includes only coding sequence for the enzyme as well as a polynucleotide which includes additional coding and/or non-coding sequence.

The present invention further relates to variants of the hereinabove described polynucleotides which encode for fragments, analogs and derivatives of the enzymes having the deduced amino acid sequences of Figures 1-10 (SEQ ID NOS:33-42). The variant of the polynucleotide may be a naturally occurring allelic variant of the polynucleotide or a non-naturally occurring variant of the polynucleotide.

Thus, the present invention includes polynucleotides encoding the same mature enzymes as shown in Figures 1-10 (SEQ ID NOS:23-32) as well as variants of such polynucleotides which variants encode for a fragment, derivative or analog of the enzymes of Figures 1-10 (SEQ ID NOS:23-32). Such nucleotide variants include deletion variants, substitution variants and addition or insertion variants.

As hereinabove indicated, the polynucleotides may have a coding sequence which is a naturally occurring allelic variant of the coding sequences shown in Figures 1-10 (SEQ ID NOS:23-32). As known in the art, an allelic variant is an alternate form of a polynucleotide sequence which may have a substitution, deletion or addition of one or more nucleotides, which does not substantially alter the function of the encoded enzyme.

Fragments of the full length gene of the present invention may be used as hybridization probes for a cDNA or a genomic library to isolate the full length DNA and to isolate other DNAs which have a high sequence similarity to the gene or similar biological activity. Probes of this type preferably have at least 10, preferably at least 15, and even more preferably at least 30 bases and may contain, for example, at least 50 or more bases. The probe may also be used to identify a DNA clone corresponding to a full length transcript and a genomic clone or clones that contain the complete gene including regulatory and promotor regions, exons and introns. An example of a screen comprises isolating the coding region of the gene by using the known DNA sequence to synthesize an oligonucleotide probe. Labeled oligonucleotides having a sequence complementary to that of the gene of the present invention are used to screen a library of genomic DNA to determine which members of the library the probe hybridizes to.

It is also appreciated that such probes can be and are preferably labeled with an analytically detectable reagent to facilitate identification of the probe. Useful reagents include but are not limited to radioactivity, fluorescent dyes or enzymes capable of catalyzing the formation of a detectable product. The probes are thus useful to isolate

complementary copies of DNA from other sources or to screen such sources for related sequences.

The present invention further relates to polynucleotides which hybridize to the hereinabove-described sequences if there is at least 70%, preferably at least 90%, and more preferably at least 95% identity between the sequences. The present invention particularly relates to polynucleotides which hybridize under stringent conditions to the hereinabove-described polynucleotides. As herein used, the term "stringent conditions" means hybridization will occur only if there is at least 95% and preferably at least 97% identity between the sequences. The polynucleotides which hybridize to the hereinabove described polynucleotides in a preferred embodiment encode enzymes which either retain substantially the same biological function or activity as the mature enzyme encoded by the DNA of Figures 1-10 (SEQ ID NOS:23-32).

Alternatively, the polynucleotide may have at least 15 bases, preferably at least 30 bases, and more preferably at least 50 bases which hybridize to any part of a polynucleotide of the present invention and which has an identity thereto, as hereinabove described, and which may or may not retain activity. For example, such polynucleotides may be employed as probes for the polynucleotides of SEQ ID NOS:23-32, for example, for recovery of the polynucleotide or as a diagnostic probe or as a PCR primer.

Thus, the present invention is directed to polynucleotides having at least a 70% identity, preferably at least 90% identity and more preferably at least a 95% identity to a polynucleotide which encodes the enzymes of SEQ ID NOS:33-42 as well as fragments thereof, which fragments have at least 15 bases, preferably at least 30 bases and most preferably at least 50 bases, which fragments are at least 90% identical, preferably at least 95% identical and most preferably at least 97% identical under stringent conditions to any portion of a polynucleotide of the present invention.

The present invention further relates to enzymes which have the deduced amino

acid sequences of Figures 1-10 (SEQ ID NOS:23-32) as well as fragments, analogs and derivatives of such enzyme.

The terms "fragment," "derivative" and "analog" when referring to the enzymes of Figures 1-10 (SEQ ID NOS:33-42) mean enzymes which retain essentially the same biological function or activity as such enzymes. Thus, an analog includes a proprotein which can be activated by cleavage of the proprotein portion to produce an active mature enzyme.

The enzymes of the present invention may be a recombinant enzyme, a natural enzyme or a synthetic enzyme, preferably a recombinant enzyme.

The fragment, derivative or analog of the enzymes of Figures 1-10 (SEQ ID NOS:33-42) may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the mature enzyme is fused with another compound, such as a compound to increase the half-life of the enzyme (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the mature enzyme, such as a leader or secretory sequence or a sequence which is employed for purification of the mature enzyme or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

The enzymes and polynucleotides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

The term "isolated" means that the material is removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a

naturally-occurring polynucleotide or enzyme present in a living animal is not isolated, but the same polynucleotide or enzyme, separated from some or all of the coexisting materials in the natural system, is isolated. Such polynucleotides could be part of a vector and/or such polynucleotides or enzymes could be part of a composition, and still be isolated in that such vector or composition is not part of its natural environment.

The enzymes of the present invention include the enzymes of SEQ ID NOS:33-42 (in particular the mature enzyme) as well as enzymes which have at least 70% similarity (preferably at least 70% identity) to the enzymes of SEQ ID NOS:33-42 and more preferably at least 90% similarity (more preferably at least 90% identity) to the enzymes of SEQ ID NOS:33-42 and still more preferably at least 95% similarity (still more preferably at least 95% identity) to the enzymes of SEQ ID NOS:33-42 and also include portions of such enzymes with such portion of the enzyme generally containing at least 30 amino acids and more preferably at least 50 amino acids.

As known in the art "similarity" between two enzymes is determined by comparing the amino acid sequence and its conserved amino acid substitutes of one enzyme to the sequence of a second enzyme.

A variant, *i.e.* a "fragment", "analog" or "derivative" polypeptide, and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions, fusions and truncations, which may be present in any combination.

Among preferred variants are those that vary from a reference by conservative amino acid substitutions. Such substitutions are those that substitute a given amino acid in a polypeptide by another amino acid of like characteristics. Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxyl residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the

aromatic residues Phe, Tyr.

Most highly preferred are variants which retain the same biological function and activity as the reference polypeptide from which it varies.

Fragments or portions of the enzymes of the present invention may be employed for producing the corresponding full-length enzyme by peptide synthesis; therefore, the fragments may be employed as intermediates for producing the full-length enzymes. Fragments or portions of the polynucleotides of the present invention may be used to synthesize full-length polynucleotides of the present invention.

The present invention also relates to vectors which include polynucleotides of the present invention, host cells which are genetically engineered with vectors of the invention and the production of enzymes of the invention by recombinant techniques.

Host cells are genetically engineered (transduced or transformed or transfected) with the vectors of this invention which may be, for example, a cloning vector or an expression vector. The vector may be, for example, in the form of a plasmid, a viral particle, a phage, *etc.* The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the genes of the present invention. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

The polynucleotides of the present invention may be employed for producing enzymes by recombinant techniques. Thus, for example, the polynucleotide may be included in any one of a variety of expression vectors for expressing an enzyme. Such vectors include chromosomal, nonchromosomal and synthetic DNA sequences, *e.g.*, derivatives of SV40; bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as

vaccinia, adenovirus, fowl pox virus, and pseudorabies. However, any other vector may be used as long as it is replicable and viable in the host.

The appropriate DNA sequence may be inserted into the vector by a variety of procedures. In general, the DNA sequence is inserted into an appropriate restriction endonuclease site(s) by procedures known in the art. Such procedures and others are deemed to be within the scope of those skilled in the art.

The DNA sequence in the expression vector is operatively linked to an appropriate expression control sequence(s) (promoter) to direct mRNA synthesis. As representative examples of such promoters, there may be mentioned: LTR or SV40 promoter, the *E. coli. lac* or *trp*, the phage lambda P<sub>L</sub> promoter and other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses.

The expression vector also contains a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression.

In addition, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in *E. coli*.

The vector containing the appropriate DNA sequence as hereinabove described, as well as an appropriate promoter or control sequence, may be employed to transform an appropriate host to permit the host to express the protein.

As representative examples of appropriate hosts, there may be mentioned: bacterial cells, such as *E. coli*, *Streptomyces*, *Bacillus subtilis*; fungal cells, such as yeast; insect cells such as *Drosophila S2* and *Spodoptera Sf9*; animal cells such as CHO, COS or Bowes melanoma; adenoviruses; plant cells, *etc.* The selection of an



appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein.

More particularly, the present invention also includes recombinant constructs comprising one or more of the sequences as broadly described above. The constructs comprise a vector, such as a plasmid or viral vector, into which a sequence of the invention has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. The following vectors are provided by way of example; Bacterial: pQE70, pQE60, pQE-9 (Qiagen), pBluescript II KS, ptrc99a, pKK223-3, pDR540, pRIT2T (Pharmacia); Eukaryotic: pXT1, pSG5 (Stratagene) pSVK3, pBPV, pMSG, pSVL, SV40 (Pharmacia). However, any other plasmid or vector may be used as long as they are replicable and viable in the host.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda P<sub>R</sub>, P<sub>L</sub> and trp. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

In a further embodiment, the present invention relates to host cells containing the above-described constructs. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-Dextran mediated transfection, or electroporation (Davis, L., Dibner, M., Battey, I., *Basic Methods in*

*Molecular Biology*, (1986)).

The constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. Alternatively, the enzymes of the invention can be synthetically produced by conventional peptide synthesizers.

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual, Second Edition*, Cold Spring Harbor, N.Y., (1989), the disclosure of which is hereby incorporated by reference.

Transcription of the DNA encoding the enzymes of the present invention by higher eukaryotes is increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act on a promoter to increase its transcription. Examples include the SV40 enhancer on the late side of the replication origin bp 100 to 270, a cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, *e.g.*, the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK),  $\alpha$ -factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence

capable of directing secretion of translated enzyme. Optionally, the heterologous sequence can encode a fusion enzyme including an N-terminal identification peptide imparting desired characteristics, *e.g.*, stabilization or simplified purification of expressed recombinant product.

Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but nonlimiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM1 (Promega Biotec, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed.

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced by appropriate means (*e.g.*, temperature shift or chemical induction) and cells are cultured for an additional period.

Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, such methods are well known to those skilled in the art.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, *Cell*, 23:175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements.

The enzyme can be recovered and purified from recombinant cell cultures by methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps.

The enzymes of the present invention may be a naturally purified product, or a product of chemical synthetic procedures, or produced by recombinant techniques from a prokaryotic or eukaryotic host (for example, by bacterial, yeast, higher plant, insect and mammalian cells in culture). Depending upon the host employed in a recombinant production procedure, the enzymes of the present invention may be glycosylated or may be non-glycosylated. Enzymes of the invention may or may not also include an initial methionine amino acid residue.

Esterases are a group of key enzymes in the metabolism of fats and are found in all organisms from microbes to mammals. In the hydrolysis reaction, an ester group is hydrolysed to an organic acid and an alcohol.

Esterases enantiomerically differentiate dicarboxylic diesters and diacetates of diols. Using the approach disclosed in a commonly assigned, copending provisional application Serial No. 60/008,316, filed on December 7, 1995 and entitled "Combinatorial Enzyme Development," the disclosure of which is incorporated herein by reference in its entirety, one could convert the enantiospecificity of the esterase. Further, the thermostable esterases are believed to have superior stability at higher temperatures and in organic solvents. Thus, they are better suited for use in rigorous production processes which require robust catalysts.

There are a number of industrial and scientific applications for esterases, such as those of the present invention, including:

- 1) Esterases are useful in the dairy industry as ripening starters for cheeses, such as the Swiss-type cheeses;
- 2) Esterases are useful in the pulp and paper industry for lignin removal from cellulose pulps, for lignin solubilization by cleaving the ester linkages between aromatic acids and lignin and between lignin and hemicelluloses, and for disruption of cell wall structure when used in combination with xylanase and other xylan-degrading enzymes in biopulping and biobleaching of pulps;
- 3) Esterases are useful in the synthesis of carbohydrate derivatives, such as sugar derivatives;
- 4) Esterases are useful, when combined with xylanases and cellulases, in the

conversion of lignocellulosic wastes to fermentable sugars for producing a variety of chemicals and fuels;

5) Esterases are useful as research reagents in studies on plant cell wall structure, particularly the nature of covalent bonds between lignin and carbohydrate polymers in the cell wall matrix;

6) Esterases are also useful as research reagents in studies on mechanisms related to disease resistance in plants and the process of organic matter decomposition; and

7) Esterases are useful in selection of plants bred for production of highly digestible animal feeds, particularly for ruminant animals.

Antibodies generated against the enzymes corresponding to a sequence of the present invention can be obtained by direct injection of the enzymes into an animal or by administering the enzymes to an animal, preferably a nonhuman. The antibody so obtained will then bind the enzymes itself. In this manner, even a sequence encoding only a fragment of the enzymes can be used to generate antibodies binding the whole native enzymes. Such antibodies can then be used to isolate the enzyme from cells expressing that enzyme.

For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler and Milstein, *Nature*, 256:495-497, 1975), the trioma technique, the human B-cell hybridoma technique (Kozbor *et al.*, *Immunology Today* 4:72, 1983), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole *et al.*, in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96, 1985).

Techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce single chain antibodies to immunogenic enzyme products of this invention. Also, transgenic mice may be used to express humanized antibodies to immunogenic enzyme products of this invention.

Antibodies generated against an enzyme of the present invention may be used in screening for similar enzymes from other organisms and samples. Such screening techniques are known in the art, for example, one such screening assay is described in Sambrook *et al.*, Molecular Cloning: A Laboratory Manual (2d Ed.), Cold Spring Harbor Laboratory, Section 12.21-12.28 (1989) which is hereby incorporated by reference in its entirety.

The present invention will be further described with reference to the following examples; however, it is to be understood that the present invention is not limited to such examples. All parts or amounts, unless otherwise specified, are by weight.

In order to facilitate understanding of the following examples certain frequently occurring methods and/or terms will be described.

"Plasmids" are designated by a lower case "p" preceded and/or followed by capital letters and/or numbers. The starting plasmids herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids in accord with published procedures. In addition, equivalent plasmids to those described are known in the art and will be apparent to the ordinarily skilled artisan.

"Digestion" of DNA refers to catalytic cleavage of the DNA with a restriction enzyme that acts only at certain sequences in the DNA. The various restriction enzymes used herein are commercially available and their reaction conditions, cofactors and other requirements were used as would be known to the ordinarily skilled artisan. For

analytical purposes, typically 1  $\mu\text{g}$  of plasmid or DNA fragment is used with about 2 units of enzyme in about 20  $\mu\text{l}$  of buffer solution. For the purpose of isolating DNA fragments for plasmid construction, typically 5 to 50  $\mu\text{g}$  of DNA are digested with 20 to 250 units of enzyme in a larger volume. Appropriate buffers and substrate amounts for particular restriction enzymes are specified by the manufacturer. Incubation times of about 1 hour at 37°C are ordinarily used, but may vary in accordance with the supplier's instructions. After digestion the reaction is electrophoresed directly on a polyacrylamide gel to isolate the desired fragment.

Size separation of the cleaved fragments is performed using 8 percent polyacrylamide gel described by Goeddel *et al.*, *Nucleic Acids Res.*, 8:4057 (1980).

"Oligonucleotides" refers to either a single stranded polydeoxynucleotide or two complementary polydeoxynucleotide strands which may be chemically synthesized. Such synthetic oligonucleotides have no 5' phosphate and thus will not ligate to another oligonucleotide without adding a phosphate with an ATP in the presence of a kinase. A synthetic oligonucleotide will ligate to a fragment that has not been dephosphorylated.

"Ligation" refers to the process of forming phosphodiester bonds between two double stranded nucleic acid fragments (Maniatis, T., *et al.*, *Id.*, p. 146). Unless otherwise provided, ligation may be accomplished using known buffers and conditions with 10 units of T4 DNA ligase ("ligase") per 0.5  $\mu\text{g}$  of approximately equimolar amounts of the DNA fragments to be ligated.

Unless otherwise stated, transformation was performed as described in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual* (2d Ed.), Cold Spring Harbor Press (1989).

### **Example 1**



### Bacterial Expression and Purification of Esterases

DNA encoding the enzymes of the present invention, SEQ ID NOS:33 through 42, were initially amplified from a pBluescript vector containing the DNA by the PCR technique using the primers noted herein. The amplified sequences were then inserted into the respective PQE vector listed beneath the primer sequences, and the enzyme was expressed according to the protocols set forth herein. The 5' and 3' primer sequences for the respective genes are as follows:

*Staphylothermus marinus* F1-12LC

5' CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGTCTTTA AACAAGCACT CT

3' CGGAAGATCT CTATCGTTTA GTGTATGATT T

vector: pQET

*Pyrodictium* TAG11-17LC

5' CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGAACTC CTTGAGCCCA CA

EcoRI

3' CGGAAGATCT CGCCGGTACA CCATCAGCCA C

BglIII

vector: pQET

*Archaeoglobus venificus* SNP6-24LC

5' CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGCCATAT GTTAGGAATG GT

3' CGGAGGTACC TTAGAACTGT GCTGAAGAAA TAAATTCGTC CATTGCTCT

3' CGGAGGTACC TTAGAACTGT GCTGAAGAAA TAAATTCGTC CATTGCTCTA TTA

vector: pQET

*Aquifex pyrophilus* - 28LC

5' CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGAGATTG AGGAAATTTG AAG

3' CGGAGGTACC CTATTCAGAA AGTACCTCTA A

vector: pQET

M11TL - 29LC

5' CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGTTTAAT ATCAATGTCT TT

3' CGGAAGATCT TTAAGGATTT TCCCTGGTA G

vector: pQET

*Thermococcus* CL-2 - 30LC

5' CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGGAGGTT TACAAGGCCA AA

3' CGGAGGTACC TTATTGAGCC GAAGAGTACG A  
vector: pQET

*Aquifex* VF5 - 34LC

5' CCGAGAATTC ATTAAAGAGG AGAAATTAACTATGATTGGC AATTGAAAT TGA EcoRI  
3' CGGAGGTACC TTAAAGTGCT CTCATATCCC C KpnI  
vector: pQET

*Teredinibacter* 42L

5' CCGAGAATTC ATTAAAGAGG AGAAATTAACTATGCCAGCT AATGACTCAC CC  
3' CGGAAGATCT TCAACAGGCT CCAATAATT TC (without His-tag)  
3' CGGAAGATCT ACAGGCTCCA AATAATTTC (with His-tag)  
vector: pQE12

*Archaeoglobus fulgidus* VC16-16MC

5' CCGAGAATTC ATTAAAGAGG AGAAATTAACTATGCTTGAT ATGCCAATCG AC EcoRI  
3' CGGAGGTACC CTAGTCGAAG ACAACAAGAG C KpnI  
vector: pQET

*Sulfolobus solfataricus* P1-8LC

5' CCGAGAATTC ATTAAAGAGG AGAAATTAACTATGCCCCAG GATCCTAGAA TT EcoRI  
3' CGGAGGTACC TTAAATTTTA TCATAAAATA C KpnI  
vector: pQET

The restriction enzyme sites indicated correspond to the restriction enzyme sites on the bacterial expression vector indicated for the respective gene (Qiagen, Inc. Chatsworth, CA). The pQE vector encodes antibiotic resistance (Amp<sup>r</sup>), a bacterial origin of replication (ori), an IPTG-regulatable promoter operator (P/O), a ribosome binding site (RBS), a 6-His tag and restriction enzyme sites.

The pQE vector was digested with the restriction enzymes indicated. The amplified sequences were ligated into the respective pQE vector and inserted in frame with the sequence encoding for the RBS. The ligation mixture was then used to transform the *E. coli* strain M15/pREP4 (Qiagen, Inc.) by electroporation. M15/pREP4 contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance

(Kan<sup>r</sup>). Transformants were identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies were selected. Plasmid DNA was isolated and confirmed by restriction analysis. Clones containing the desired constructs were grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture was used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells were grown to an optical density 600 (O.D.<sup>600</sup>) of between 0.4 and 0.6. IPTG ("Isopropyl-B-D-thiogalacto pyranoside") was then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression. Cells were grown an extra 3 to 4 hours. Cells were then harvested by centrifugation.

The primer sequences set out above may also be employed to isolate the target gene from the deposited material by hybridization techniques described above.

### **Example 2**

#### **Isolation of a Selected Clone from the Deposited Genomic Clones**

The two oligonucleotide primers corresponding to the gene of interest are used to amplify the gene from the deposited material. A polymerase chain reaction is carried out in 25 µl of reaction mixture with 0.1 µg of the DNA of the gene of interest. The reaction mixture is 1.5-5 mM MgCl<sub>2</sub>, 0.01 % (w/v) gelatin, 20 µM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 1.25 Unit of Taq polymerase. Thirty cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with the Perkin-Elmer Cetus 9600 thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the gene of interest by subcloning and sequencing the DNA product.

### Example 3

#### Production of the Expression Gene Bank

Colonies containing pBluescript plasmids with random inserts from the organisms M11TL, *Thermococcus* GU5L5, and *Teredinibacter* were obtained according to the method of Hay and Short, *Strategies*, 5:16, 1992.

### Example 4

#### Screening for Lipase/Esterase Activity

The resulting colonies were picked with sterile toothpicks and used to singly inoculate each of the wells of 96-well microtiter plates. The wells contained 250  $\mu$ L of LB media with 100  $\mu$ g/mL ampicillin, 80  $\mu$ g/mL methicillin, and 10% v/v glycerol (LB Amp/Meth, glycerol). The cells were grown overnight at 37°C without shaking. This constituted generation of the "Source GeneBank." Each well of the Source GeneBank thus contained a stock culture of *E. coli* cells, each of which contained a pBluescript with a unique DNA insert.

The plates of the Source GeneBank were used to multiply inoculate a single plate (the "Condensed Plate") containing in each well 200  $\mu$ L of LB Amp/Meth, glycerol. This step was performed using the High Density Replicating Tool (HDRT) of the Beckman Biomek with a 1% bleach, water, isopropanol, air-dry sterilization cycle in between each inoculation. Each well of the Condensed Plate thus contained 10 to 12 different pBluescript clones from each of the source library plates. The Condensed Plate was grown for 16 hours at 37°C and then used to inoculate two white 96-well Polyfiltronics microtiter daughter plates containing in each well 250  $\mu$ L of LB Amp/Meth (no glycerol). The original condensed plate was put in storage -80°C. The two condensed daughter plates were incubated at 37°C for 18 hours.

The short chain esterase '600  $\mu$ M substrate stock solution' was prepared as follows:

25 mg of each of the following compounds was dissolved in the appropriate volume of DMSO to yield a 25.2 mM solution. The compounds used were 4-methylumbelliferyl propionate, 4-methylumbelliferyl butyrate, and 4-methylumbelliferyl heptanoate. Two hundred fifty microliters of each DMSO solution was added to ca 9 mL of 50 mM, pH 7.5 Hepes buffer which contained 0.6% of Triton X-100 and 0.6 mg per mL of dodecyl maltoside (Anatrace). The volume was taken to 10.5 mL with the above Hepes buffer to yield a slightly cloudy suspension.

The long chain '600  $\mu$ M substrate stock solution' was prepared as follows: 25 mg of each of the following compounds was dissolved in DMSO to 25.2 mM as above. The compounds used were 4-methylumbelliferyl elaidate, 4-methylumbelliferyl palmitate, 4-methylumbelliferyl oleate, and 4-methylumbelliferyl stearate. All required brief warming in a 70°C bath to achieve dissolution. Two hundred fifty microliters of each DMSO solution was added to the Hepes buffer and diluted to 10.5 mL as above. All seven umbelliferones were obtained from Sigma Chemical Co.

Fifty  $\mu$ L of the long chain esterase or short chain esterase '600  $\mu$ M substrate stock solution' was added to each of the wells of a white condensed plate using the Biomek to yield a final concentration of substrate of about 100  $\mu$ M.. The fluorescence values were recorded (excitation = 326 nm, emission = 450 nm) on a plate-reading fluorometer immediately after addition of the substrate. The plate was incubated at 70°C for 60 minutes in the case of the long chain substrates, and 30 minutes at RT in the case of the short chain substrates. The fluorescence values were recorded again. The initial and final fluorescence values were compared to determine if an active clone was present.

### **Example 5**

#### **Isolation and Purification of the Active Clone**

To isolate the individual clone which carried the activity, the Source GeneBank plates were thawed and the individual wells used to singly inoculate a new plate containing

LB Amp/Meth. As above, the plate was incubated at 37°C to grow the cells, 50 µL of 600 µM substrate stock solution was added using the Biomek and the fluorescence was determined. Once the active well from the source plate was identified, cells from this active well were streaked on agar with LB/Amp/Meth and grown overnight at 37°C to obtain single colonies. Eight single colonies were picked with a sterile toothpick and used to singly inoculate the wells of a 96-well microtiter plate. The wells contained 250 µL of LB Amp/Meth. The cells were grown overnight at 37°C without shaking. A 200 µL aliquot was removed from each well and assayed with the appropriate long or short chain substrates as above. The most active clone was identified and the remaining 50 µL of culture was used to streak an agar plate with LB/Amp/Meth. Eight single colonies were picked, grown and assayed as above. The most active clone was used to inoculate 3 mL cultures of LB/Amp/Meth, which were grown overnight. The plasmid DNA was isolated from the cultures and utilized for sequencing.

Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, within the scope of the appended claims, the invention may be practiced otherwise than as particularly described.

## SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANTS:

ROBERTSON, Daniel E.

MURPHY, Dennis

REID, John

MAFFIA, Anthony

LINK, Steven

SWANSON, Ronald V.

WARREN, Patrick V.

KOSMOTKA, Anna

(ii) TITLE OF INVENTION:

ESTERASES

(iii) NUMBER OF SEQUENCES: 42

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
CECCHI, STEWART & OLSTEIN  
(B) STREET: 6 BECKER FARM ROAD  
(C) CITY: ROSELAND  
(D) STATE: NEW JERSEY  
(E) COUNTRY: USA  
(F) ZIP: 07068

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: 3.5 INCH DISKETTE  
(B) COMPUTER: IBM PS/2  
(C) OPERATING SYSTEM: MS-DOS  
(D) SOFTWARE: WORD PERFECT 5.1

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: Unassigned  
(B) FILING DATE: Concurrently  
(C) CLASSIFICATION: Unassigned

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: HERRON, CHARLES J.  
(B) REGISTRATION NUMBER: 28,019  
(C) REFERENCE/DOCKET NUMBER: 331400-39

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 201-994-1700  
(B) TELEFAX: 201-994-1744

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS  
(A) LENGTH: 52 NUCLEOTIDES  
(B) TYPE: NUCLEIC ACID  
(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGTCTTTA AACAAGCACT CT

52

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS  
(A) LENGTH: 31 NUCLEOTIDES  
(B) TYPE: NUCLEIC ACID  
(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CGGAAGATCT CTATCGTTTA GTGTATGATT T

31

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS  
(A) LENGTH: 52 NUCLEOTIDES  
(B) TYPE: NUCLEIC ACID  
(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGAACTC CTTGAGCCCA CA

52

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS  
(A) LENGTH: 31 NUCLEOTIDES  
(B) TYPE: NUCLEIC ACID  
(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CGGAAGATCT CGCCGGTACA CCATCAGCCA C

31



(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS
  - (A) LENGTH: 52 NUCLEOTIDES
  - (B) TYPE: NUCLEIC ACID
  - (C) STRANDEDNESS: SINGLE
  - (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CCGAGAATTC ATTAAAGAGG AGAAATTAC TATGCCATAT GTTAGGAATG GT

52

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS
  - (A) LENGTH: 53 NUCLEOTIDES
  - (B) TYPE: NUCLEIC ACID
  - (C) STRANDEDNESS: SINGLE
  - (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CGGAGGTACC TTAGAACTGT GCTGAAGAA TAAATTCGTC CATTGCTCTA TTA

53

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS
  - (A) LENGTH: 49 NUCLEOTIDES
  - (B) TYPE: NUCLEIC ACID
  - (C) STRANDEDNESS: SINGLE
  - (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGGAGGTACC TTAGAACTGT GCTGAAGAA TAAATTCGTC CATTGCTCT

49

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS
  - (A) LENGTH: 53 NUCLEOTIDES
  - (B) TYPE: NUCLEIC ACID
  - (C) STRANDEDNESS: SINGLE
  - (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CCGAGAATTC ATTAAAGAGG AGAAATTAC TATGAGATTG AGGAAATTTG AAG

53

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS
  - (A) LENGTH: 31 NUCLEOTIDES
  - (B) TYPE: NUCLEIC ACID
  - (C) STRANDEDNESS: SINGLE
  - (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CGGAGGTACC CTATTCAGAA AGTACCTCTA A

31

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS
  - (A) LENGTH: 52 NUCLEOTIDES
  - (B) TYPE: NUCLEIC ACID
  - (C) STRANDEDNESS: SINGLE
  - (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGTTTAAT ATCAATGTCT TT

52

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS
  - (A) LENGTH: 31 NUCLEOTIDES
  - (B) TYPE: NUCLEIC ACID
  - (C) STRANDEDNESS: SINGLE
  - (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CGGAAGATCT TTAAGGATT TCCCTGGCTA G

31

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS
  - (A) LENGTH: 52 NUCLEOTIDES
  - (B) TYPE: NUCLEIC ACID
  - (C) STRANDEDNESS: SINGLE
  - (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGGAGGTT TACAAGGCCA AA

52

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS
  - (A) LENGTH: 31 NUCLEOTIDES
  - (B) TYPE: NUCLEIC ACID

(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGGAGGTACC TTATTGAGCC GAAGAGTACG A

31

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS  
(A) LENGTH: 53 NUCLEOTIDES  
(B) TYPE: NUCLEIC ACID  
(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CCGAGAATTC ATTAAAGAGG AGAAATTAC TATGATTGGC AATTGAAAT TGA

53

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS  
(A) LENGTH: 31 NUCLEOTIDES  
(B) TYPE: NUCLEIC ACID  
(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CGGAGGTACC TTAAAGTGCT CTCATATCCC C

31

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS  
(A) LENGTH: 31 NUCLEOTIDES  
(B) TYPE: NUCLEIC ACID  
(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CCGAGAATTC ATTAAAGAGG AGAAATTAC TATGCCAGCT AATGACTCAC CC

52

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS  
(A) LENGTH: 32 NUCLEOTIDES  
(B) TYPE: NUCLEIC ACID  
(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:  
CGGAAGATCT TCAACAGGCT CCAAATAATT TC 32

(2) INFORMATION FOR SEQ ID NO:18:  
(i) SEQUENCE CHARACTERISTICS  
(A) LENGTH: 29 NUCLEOTIDES  
(B) TYPE: NUCLEIC ACID  
(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: LINEAR  
  
(ii) MOLECULE TYPE: cDNA  
  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:  
CGGAAGATCT ACAGGCTCCA AATAATTIC 29

(2) INFORMATION FOR SEQ ID NO:19:  
(i) SEQUENCE CHARACTERISTICS  
(A) LENGTH: 52 NUCLEOTIDES  
(B) TYPE: NUCLEIC ACID  
(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: LINEAR  
  
(ii) MOLECULE TYPE: cDNA  
  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:  
CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGCTTGAT ATGCCAATCG AC 52

(2) INFORMATION FOR SEQ ID NO:20:  
(i) SEQUENCE CHARACTERISTICS  
(A) LENGTH: 31 NUCLEOTIDES  
(B) TYPE: NUCLEIC ACID  
(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: LINEAR  
  
(ii) MOLECULE TYPE: cDNA  
  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:  
CGGAGGTACC CTAGTCGAAC AGAAGAACAG C 31

(2) INFORMATION FOR SEQ ID NO:21:  
(i) SEQUENCE CHARACTERISTICS  
(A) LENGTH: 52 NUCLEOTIDES  
(B) TYPE: NUCLEIC ACID  
(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: LINEAR  
  
(ii) MOLECULE TYPE: cDNA  
  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:  
CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGCCCTA GATCCTAGAA TT 52

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS  
(A) LENGTH: 31 NUCLEOTIDES  
(B) TYPE: NUCLEIC ACID  
(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CGGAGGTACC TTAAATTTTA TCATAAAATA C

31

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS  
(A) LENGTH: 555 NUCLEOTIDES  
(B) TYPE: NUCLEIC ACID  
(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

|   |     |
|---|-----|
| ATG TCT TTA AAC AAG CAC TCT TCG ATG GAT ATG ATA ATA TTT ATT CTC | 48  |
| Met Ser Leu Asn Lys His Ser Trp Met Asp Met Ile Ile Phe Ile Leu |     |
| 1 5 10 15   |     |
| AGC TTT TCT TTC CCA TTA ACA ATG ATC GCA TTA GCT ATC TCT ATG TCG | 96  |
| Ser Phe Ser Phe Pro Leu Thr Met Ile Ala Leu Ala Ile Ser Met Ser |     |
| 20 25 30  |     |
| TCA TGG TTT AAT ATA TGG AAT AAT GCA TTA AGC GAT CTA GGA CAT GCT | 144 |
| Ser Trp Phe Asn Ile Trp Asn Asn Ala Leu Ser Asp Leu Gly His Ala |     |
| 35 40 45  |     |
| GTT AAA AGC AGT GTT GCT CCA ATA TTC AAT CTA GGT CTT GCA ATT GGT | 192 |
| Val Lys Ser Ser Val Ala Pro Ile Phe Asn Leu Gly Leu Ala Ile Gly |     |
| 50 55 60  |     |
| GGG ATA CTA ATT GTT ATA GTT GGT TTA AGA AAT CTT TAT TCG TGG AGT | 240 |
| Gly Ile Leu Ile Val Ile Val Gly Leu Arg Asn Leu Tyr Ser Trp Ser |     |
| 65 70 75 80   |     |
| AGA GTT AAA GGA TCT TTA ATC ATA TCC ATG GGT GTA TTT CTT AAC TTA | 288 |
| Arg Val Lys Gly Ser Leu Ile Ile Ser Met Gly Val Phe Leu Asn Leu |     |
| 85 90 95  |     |
| ATA GGG GTT TTC GAC GAA GTA TAT GGT TGG ATA CAT TTC CTA GTC TCA | 336 |
| Ile Gly Val Phe Asp Glu Val Tyr Gly Trp Ile His Phe Leu Val Ser |     |
| 100 105 110   |     |
| GTA TTG TTT TTC TTA TCA ATA ATA GCA TAT TTC ATA GCT ATA TCA ATA | 384 |
| Val Leu Phe Phe Leu Ser Ile Ile Ala Tyr Phe Ile Ala Ile Ser Ile |     |
| 115 120 125   |     |
| CTT GAC AAA TCA TGG ATA GCT GTT CTA CTA ATA ATA GGT CAT ATT GCA | 432 |
| Leu Asp Lys Ser Trp Ile Ala Val Leu Leu Ile Ile Gly His Ile Ala |     |
| 130 135 140   |     |
| ATG TGG TAT CTA CAC TTT GCT TCA GAG ATT CCG AGA GGT GCT GCT ATT | 480 |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Trp | Tyr | Leu | His | Phe | Ala | Ser | Glu | Ile | Pro | Arg | Gly | Ala | Ala | Ile |     |
| 145 |     |     |     |     | 150 |     |     |     |     | 155 |     |     |     |     | 160 |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| CCC | GAG | TTA | TTA | GCG | GTA | TTC | TCG | TTT | TTA | CCA | TTC | TAT | ATA | AGA | CAG | 528 |
| Pro | Glu | Leu | Leu | Ala | Val | Phe | Ser | Phe | Leu | Pro | Phe | Tyr | Ile | Arg | Asp |     |
|     |     |     |     | 165 |     |     |     |     | 170 |     |     |     |     | 175 |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| TAT | TTT | AAA | TCA | TAC | ACT | AAA | CGA | TAG |     |     |     |     |     |     |     | 576 |
| Tyr | Phe | Lys | Ser | Tyr | Thr | Lys | Arg |     |     |     |     |     |     |     |     |     |
|     |     |     |     | 180 |     |     |     |     |     |     |     |     |     |     |     |     |

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 1041 NUCLEOTIDES

(B) TYPE: NUCLEIC ACID

(C) STRANDEDNESS: SINGLE

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ATG | AAA | CTC | CTT | GAG | CCC | ACA | AAT | ACC | TCC | TAC | ACG | CTG | TTA | CAG | GAT | 48  |
| Met | Lys | Leu | Leu | Glu | Pro | Thr | Asn | Thr | Ser | Tyr | Thr | Leu | Leu | Gln | Asp |     |
| 1   |     |     |     | 5   |     |     |     | 10  |     |     |     |     |     | 15  |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| TTA | GCA | TTG | CAT | TTT | GCA | TTT | TAC | TGG | TTT | CTG | GCC | GTG | TAT | ACG | TGG | 96  |
| Leu | Ala | Leu | His | Phe | Ala | Phe | Tyr | Trp | Phe | Leu | Ala | Val | Tyr | Thr | Trp |     |
|     |     |     | 20  |     |     |     | 25  |     |     |     |     |     | 30  |     |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| TTA | CCC | GGT | GTC | CTA | GTC | CGG | GCC | GTA | GCT | GTG | GAC | ACA | GGG | GTG | GCT | 144 |
| Leu | Pro | Gly | Val | Leu | Val | Arg | Gly | Val | Ala | Val | Asp | Thr | Gly | Val | Ala |     |
|     |     | 35  |     |     |     | 40  |     |     |     |     | 45  |     |     |     |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| CGG | GTG | CCT | GGG | CTC | GGC | CGG | CCC | GGT | AAG | AGG | CTG | CTC | CTG | GCC | GCT | 192 |
| Arg | Val | Pro | Gly | Leu | Gly | Arg | Arg | Gly | Lys | Arg | Leu | Leu | Leu | Ala | Ala |     |
|     | 50  |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| GTG | GCT | GTC | TTG | GCG | CTT | GTT | GTG | TCC | GTT | GTT | GTC | CCG | GCT | TAT | GTG | 240 |
| Val | Ala | Val | Leu | Ala | Leu | Val | Val | Ser | Val | Val | Val | Pro | Ala | Tyr | Val |     |
|     | 65  |     |     | 70  |     |     | 75  |     |     |     |     |     |     | 80  |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| GCG | TAT | AGT | AGT | CTG | CAC | CCG | GAG | AGC | TGT | CGG | CCC | GTT | GCG | CCG | GAG | 288 |
| Ala | Tyr | Ser | Ser | Leu | His | Pro | Glu | Ser | Cys | Arg | Pro | Val | Ala | Pro | Glu |     |
|     |     |     |     | 85  |     |     | 90  |     |     |     |     |     |     | 95  |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| GGG | CTC | ACC | TAC | AAA | GAG | TTC | ACC | GTG | ACC | GCG | GAG | GAT | GGC | TTG | GTG | 336 |
| Gly | Leu | Thr | Tyr | Lys | Glu | Phe | Ser | Val | Thr | Ala | Glu | Asp | Gly | Leu | Val |     |
|     |     |     | 100 |     |     |     | 105 |     |     |     |     |     | 110 |     |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| GTT | CGG | GGC | TGG | GTG | CTG | GGC | CCC | GGC | GCT | GGG | GGC | AAC | CCG | GTG | TTC | 384 |
| Val | Arg | Gly | Trp | Cal | Leu | Gly | Pro | Gly | Ala | Gly | Gly | Asn | Pro | Val | Phe |     |
|     |     | 115 |     |     |     | 120 |     |     |     |     |     | 125 |     |     |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| GTT | TTG | ATG | CAC | GGG | TAT | ACT | GGG | TGC | CGC | TCG | GCG | CCC | TAC | ATG | GCT | 432 |
| Val | Leu | Met | His | Gly | Tyr | Thr | Gly | Cys | Arg | Ser | Ala | Pro | Tyr | Met | Ala |     |
|     | 130 |     |     |     | 135 |     |     |     |     | 140 |     |     |     |     |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| GTG | CTG | GCC | CGG | GAG | CTC | GTG | GAG | TGG | GGG | TAC | CCG | GTG | GTT | GTG | TTC | 480 |
| Val | Leu | Ala | Arg | Glu | Leu | Val | Glu | Trp | Gly | Tyr | Pro | Val | Val | Val | Phe |     |
| 145 |     |     |     |     | 150 |     |     |     | 155 |     |     |     |     |     | 160 |     |

|   |      |
|---|------|
| GAC TTC CGG GGC CAC GGG GAG AGC GGG GGC TCG ACG ACG ATT GGG CCC<br>Asp Phe Arg Gly His Gly Glu Ser Gly Gly Ser Thr Thr Ile Gly Pro<br>165 170 175     | 528  |
| CGG GAG GTG CTG GAT GCC CGG GCT GTG GTG GGC TAT GTC TCG GAG CGG<br>Arg Glu Val Leu Asp Ala Arg Ala Val Val Gly Tyr Val Ser Glu Arg<br>180 185 190     | 576  |
| TTC CCC GGC CGC CGG ATA ATA TTG GTG GGG TTC AGT ATG GGC GGC GCT<br>Phe Pro Gly Arg Arg Ile Ile Leu Val Gly Phe Ser Met Gly Gly Ala<br>195 200 205     | 624  |
| GTA GCG ATC GTG GAG GGT GCT GGG GAC CCG CGG GTC TAC GCG GTG GCT<br>Val Ala Ile Val Glu Gly Ala Gly Asp Pro Arg Val Tyr Ala Val Ala<br>210 215 220     | 672  |
| GCT GAT AGC CCG TAC TAT AGG CTC CGG GAC GTC ATA CCC CGG TGG CTG<br>Ala Asp Ser Pro Tyr Tyr Arg Leu Arg Asp Val Ile Pro Arg Trp Leu<br>225 230 235 240 | 720  |
| GAG TAC AAG ACG CCG CTG CCG GGC TGG GTG GGT GTG CTG GCC GGG TTC<br>Glu Tyr Lys Thr Leu Pro Gly Trp Val Gly Val Leu Ala Gly Phe<br>245 250 255         | 768  |
| TAC GGG AGG CTG ATG GCG GGC GTT GAC CTC GGC TTC GGC CCC GCT GGG<br>Tyr Gly Arg Leu Met Ala Gly Val Asp Leu Gly Phe Gly Pro Ala Gly<br>260 265 270     | 816  |
| GTG GAG CGC GTG GAT AAG CCG TTG CTG GTG GTG TAT GGG CCC CGG GAC<br>Val Gly Arg Val Asp Lys Pro Leu Leu Val Val Tyr Gly Pro Arg Asp<br>275 280 285     | 864  |
| CCG CTG GTG ACG CGG GAC GAG GCG AGG AGC CTG GCG TCC CGT AGC CCG<br>Pro Leu Val Thr Arg Asp Glu Ala Arg Ser Leu Ala Ser Arg Ser Pro<br>290 295 300     | 912  |
| TGT GGC CGT CTC GTC GAG GTT CCT GGG GCT GGC CAC GTG GAG GCC GTG<br>Cys Gly Arg Leu Val Glu Val Pro Gly Ala Gly His Val Glu Ala Val<br>305 310 315 320 | 960  |
| GAT GTG CTC GGG CCG GGC CGC TAC GCA GAC ATG CTG ATA GAG CTG GCG<br>Asp Val Leu Gly Pro Gly Arg Tyr Ala Asp Met Leu Ile Glu Leu Ala<br>325 330 335     | 1008 |
| CAC GAG GAG TGC CCT CCG GGG GGC GGT GGC TGA<br>His Glu Glu Cys Pro Pro Gly Ala Gly Gly<br>340 345   | 1019 |

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 789 NUCLEOTIDES
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

|   |     |
|---|-----|
| ATG CCA TAT GTT AGG AAT GGT GGT GTA AAT ATC TAT TAT GAA CTG GTG<br>Met Pro Tyr Val Arg Asn Gly Gly Val Asn Ile Tyr Tyr Glu Leu Val<br>1 5 10 15       | 48  |
| GAT GGA CCT GAG CCA CCA ATT GTC TTT GTT CAC GGA TGG ACA GCA AAT<br>Asp Gly Pro Glu Pro Pro Ile Val Phe Val His Gly Trp Thr Ala Asn<br>20 25 30        | 96  |
| ATG AAT TTT TGG AAA GAG CAA ACA CGT TAT TTT GCA GGC AGG AAT ATG<br>Met Asn Phe Trp Lys Glu Gln Arg Arg Tyr Phe Ala Gly Arg Asn Met<br>35 40 45        | 144 |
| ATG TTG TTT GTC GAT AAC AGA GGT CAT GGC AGG TCC GAT AAG CCA CTT<br>Met Leu Phe Val Asp Asn Arg Gly His Gly Arg Ser Asp Lys Pro Leu<br>50 55 60        | 192 |
| GGA TAC GAT TTC TAC AGA TTT GAG AAC TTC ATT TCA GAT TTA GAT GCG<br>Gly Tyr Asp Phe Tyr Arg Phe Glu Asn Phe Ile Ser Asp Leu Asp Ala<br>65 70 75 80     | 240 |
| GTT GTT AGG GAG ACT GGA GTG GAG AAA TTT GTT CTC GTC GGA CAT TCA<br>Val Val Arg Glu Thr Gly Val Glu Lys Phe Cal Leu Val Gly His Ser<br>85 90 95        | 288 |
| TTC GGA ACA ATG ATC TCT ATG AAG TAC TGT TCG GAG TAT CGG AAT CGG<br>Phe Gly Thr Met Ile Ser Met Lys Tyr Cys Ser Glu Tyr Arg Asn Arg<br>100 105 110     | 336 |
| GTT CTT GCT CTA ATC CTC ATA GGT GGT GGC AGC AGA ATA AAG CTT CTA<br>Val Leu Ala Leu Ile Leu Ile Gly Gly Ser Arg Ile Lys Leu Leu<br>115 120 125         | 384 |
| CAC AGA ATT GGA TAT CCT TTA GCA AAG ATT CTT GCA TCC ATT GCA TAC<br>His Arg Ile Gly Tyr Pro Leu Ala Lys Ile Leu Ala Ser Ile Ala Tyr<br>130 135 140     | 432 |
| AAG AAG TCT TCA AGA TTG GTC GCA GAT CTT TCC TTT GGC AAA AAT GCT<br>Lys Lys Ser Ser Arg Leu Val Ala Asp Leu Ser Phe Gly Lys Asn Ala<br>145 150 155 160 | 480 |
| GGT GAA CTT AAA GAG TGG GGA TGG AAA CAG GCA ATG GAT TAT ACA CCC<br>Gly Glu Leu Lys Glu Trp Gly Trp Lys Gln Ala Met Asp Tyr Thr Pro<br>165 170 175     | 528 |
| TCC TAC GTG GCA ATG GAC ACG TAC AGA ACT CTA ACG AAA GTG AAT CTT<br>Ser Tyr Val Ala Met Tyr Thr Tyr Arg Thr Leu Thr Lys Val Asn Leu<br>180 185 190     | 576 |
| GAA AAT ATC TTG GAG AAA ATA GAC TGT CCA ACA CTG ATT ATC GTT GGA<br>Glu Asn Ile Leu Glu Lys Ile Asp Cys Pro Thr Leu Ile Ile Val Gly<br>195 200 205     | 624 |
| GAA GAG GAT GCA CTA TTG CCC GTT AGC AAA TCA GTT GAG CTG AGC AGG<br>Glu Glu Asp Ala Leu Leu Pro Val Ser Lys Ser Val Glu Leu Ser Arg<br>210 215 220     | 672 |
| AGG ATA GAA AAC TCA AAG CTT GTG ATC ATC CCA AAC TCG GGG CAT TGC<br>Arg Ile Glu Asn Ser Lys Leu Val Ile Ile Pro Asn Ser Gly His Cys<br>225 230 235 240 | 720 |
| GTA ATG CTT GAG AGT CCA AGT GAG GTT AAT AGA GCA ATG GAC GAA TTC<br>Val Met Leu Glu Ser Pro Ser Glu Val Asn Arg Ala Met Asp Glu Phe<br>245 250 255     | 768 |



ATT TCT TCA GCA CAG TTC TAA  
 Ile Ser Ser Ala Gln Phe  
 260

774

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 756 NUCLEOTIDES
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

|   |     |
|---|-----|
| TTG AGA TTG AGG AAA TTT GAA GAG ATA AAC CTC GTT CTT TCG GGA GGA<br>Leu Arg Leu Arg Lys Phe Glu Glu Ile Asn Leu Val Leu Ser Gly Gly<br>1 5 10 15       | 48  |
| GCT GCA AAG GGC ATA GCC CAC ATA GGT GTT TTG AAA GCT ATA AAC GAG<br>Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys Ala Ile Asn Glu<br>20 25 30        | 96  |
| CTC GGT ATA AGG GTG AGG GCT TTA AGC GGG GTG AGC GCC GGG GCA ATC<br>Leu Glu Ile Arg Val Arg Ala Leu Ser Gly Val Ser Ala Gly Ala Ile<br>35 40 45        | 144 |
| GTT TCG GTC TTT TAT GCC TCA GGC TAC TCC CCT GAA GGG ATG TTC AGC<br>Val Ser Val Phe Tyr Ala Ser Gly Tyr Ser Pro Glu Gly Met Phe Ser<br>50 55 60        | 192 |
| CTT CTG AAG AGG GTA AAC TGG CTG AAG CTG TTT AAG TTC AAG CCA CCT<br>Leu Leu Lys Arg Val Asn Trp Leu Lys Leu Lys Phe Lys Pro Pro<br>65 70 75 80         | 240 |
| CTG AAG GGA TTG ATA GGG TGG GAG AAG GCT ATA AGA TTC CTT GAG GAA<br>Leu Lys Gly Leu Ile Gly Trp Glu Lys Ala Ile Arg Phe Leu Glu Glu<br>85 90 95        | 288 |
| GTT CTC CCT TAC AGG AGA ATA GAA AAA CTT GAG ATA CCG ACG TAT ATA<br>Val Leu Pro Tyr Arg Arg Ile Glu Lys Leu Glu Ile Pro Thr Tyr Ile<br>100 105 110     | 336 |
| TGC GCG ACG GAT TTA TAC TCG GCA AGG GCT CTA TAC CTC TCG GAA GGG<br>Cys Ala Thr Asp Leu Tyr Ser Gly Arg Ala Leu Tyr Leu Ser Glu Gly<br>115 120 125     | 384 |
| AGT TTA ATC CCC GCA CTT CTC GGC AGC TGT GCA ATT CCC GGC ATA TTT<br>Ser Leu Ile Pro Ala Leu Leu Gly Ser Cys Ala Ile Pro Gly Ile Phe<br>130 135 140     | 432 |
| GAA CCC GTT GAG TAT AAG AAT TAC TTG CTC GTT GAC GGA GGT ATA GTT<br>Glu Pro Val Glu Tyr Lys Asn Tyr Leu Leu Val Asp Gly Gly Ile Val<br>145 150 155 160 | 480 |
| AAC AAC CTT CCC GTT GAG CCC TTT CAG GAA AGC GGT ATT CCC ACC GTT<br>Asn Asn Leu Pro Val Glu Pro Phe Gln Glu Ser Gly Ile Pro Thr Val<br>165 170 175     | 528 |
| TGC GTT GAT GTC CTT CCC ATA GAG CCG GAA AAG GAT ATA AAG AAC ATT<br>Cys Val Asp Val Leu Pro Ile Glu Pro Glu Lys Asp Ile Lys Asn Ile<br>180 185 190     | 576 |

|   |     |
|---|-----|
| CTT CAC ATC CTT TTG AGG AGC TTC TTT CTT GCG GTC CGC TCA AAC TCC | 624 |
| Leu His Ile Leu Leu Arg Ser Phe Phe Leu Ala Val Arg Ser Asn Ser |     |
| 195 200 205   |     |
| GAA AAG AGA AAG GAG TTT TGT GAC CTC GTT ATA GTT CCT GAG CTT GAG | 672 |
| Glu Lys Arg Lys Glu Phe Cys Asp Leu Val Ile Val Pro Glu Leu Glu |     |
| 210 215 220   |     |
| GAG TTC ACA CCC CTT GAT GTT ACA AAA GCG GAC CAA ATA ATG GAG AGG | 720 |
| Glu Phe Thr Pro Leu Asp Val Arg Lys Ala Asp Gln Ile Met Glu Arg |     |
| 225 230 235 240   |     |
| GGA TAC ATA AAG GCC TTA GAG TCA CTT TCT GAA TAG                 | 768 |
| Gly Tyr Ile Lys Ala Leu Glu Val Leu Ser Glu                     |     |
| 245 250   |     |

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 894 NUCLEOTIDES

(B) TYPE: NUCLEIC ACID

(C) STRANDEDNESS: SINGLE

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

|   |     |
|---|-----|
| ATG TTT AAT ATC AAT GTC TTT GTT AAT ATA TCT TGG CTG TAT TTT TCA | 48  |
| Met Phe Asn Ile Asn Val Phe Val Asn Ile Ser Trp Leu Tyr Phe Ser |     |
| 1 5 10 15   |     |
| GGG ATA GTT ATG AAG ACT GTG GAA GAG TAT GCG CTA CTT GAA ACA GGC | 96  |
| Gly Ile Val Met Lys Thr Val Glu Glu Tyr Ala Leu Leu Glu Thr Gly |     |
| 20 25 30  |     |
| GTA AGA GTG TTT TAT CGG TGT GTA ATC CCG GAG AAA GCT TTT AAC ACT | 144 |
| Val Arg Val Phe Tyr Arg Cys Val Ile Pro Glu Lys Ala Phe Asn Thr |     |
| 35 40 45  |     |
| TTG ATA ATA GGT TCA CAC GGA TTG GGG GCG CAC AGT GGA ATC TAC ATT | 192 |
| Leu Ile Ile Gly Ser His Gly Leu Gly Ala His Ser Gly Ile Tyr Ile |     |
| 50 55 60  |     |
| AGT GTT GCT GAA GAA TTT GCT AGG CAC GGA TTT GGA TTC TGC ATG CAC | 240 |
| Ser Val Ala Glu Glu Phe Ala Arg His Gly Phe Gly Phe Cys Met His |     |
| 65 70 75 80   |     |
| GAT CAA AGG GGA CAT GGG AGA ACG GCA AGC GAT AGA GAA AGA GGG TAT | 288 |
| Asp Gln Arg Gly His Gly Arg Thr Ala Ser Asp Arg Glu Arg Gly Tyr |     |
| 85 90 95  |     |
| GTG GAG GGC TTT CAC AAC TTC ATA GAG GAT ATG AAG GCC TTC TCC GAT | 336 |
| Val Glu Gly Phe His Asn Phe Ile Glu Asp Met Lys Ala Phe Ser Asp |     |
| 100 105 110   |     |
| TAT GCC AAG TGG CGC GTG GGA GGT GAC GAA ATA ATA TTG CTA GGA CAC | 384 |
| Tyr Ala Lys Trp Arg Val Gly Gly Asp Glu Ile Ile Leu Leu Gly His |     |
| 115 120 125   |     |
| AGT ATG GGC GGG CTG ATA GCG CAC GGA ACA GTT GCA ACT TAT AAA GAA | 432 |
| Ser Met Gly Gly Leu Ile Ala Leu Leu Thr Val Ala Thr Tyr Lys Glu |     |

| 130   | 135 | 140 |     |
|---|-----|-----|-----|
| ATC GCC AAG GGA GTT ATC GCG CTA GCC CCG GCC CTC CAA ATC CCC TTA<br>Ile Ala Lys Gly Val Ile Ala Leu Ala Pro Ala Leu Gln Ile Pro Leu<br>145 150 155 160 |     |     | 480 |
| ACC CCG GCT AGA AGA CTT GTT CTA AGC CTC GCG TCA AGG CTT GCC CCG<br>Thr Pro Ala Arg Arg Leu Val Leu Ser Leu Ala Ser Arg Leu Ala Pro<br>165 170 175     |     |     | 528 |
| CAT TCT AAG ATC ACC TTA CAA ACG AGA TTG CCG CAG AAA CCA GAG GGT<br>His Ser Lys Ile Thr Leu Gln Arg Arg Leu Pro Gln Lys Pro Glu Gly<br>180 185 190     |     |     | 576 |
| TTT CAA AGA GCA AAA GAT ATA GAA TAC AGT CTG AGT GAA ATA TCA GTC<br>Phe Gln Arg Ala Lys Asp Ile Glu Tyr Ser Leu Ser Glu Ile Ser Val<br>195 200 205     |     |     | 624 |
| AAG CTC GTG GAC GAA ATG ATT AAA GCA TCA TCT ATG TCT TGG ACC ATA<br>Lys Leu Val Asp Glu Met Ile Lys Ala Ser Ser Met Phe Trp Thr Ile<br>210 215 220     |     |     | 672 |
| GCA GGG GAA ATT AAT ACT CCC GTC CTG CTT ATT CAT GGG GAA AAA CAG<br>Ala Gly Glu Ile Asn Thr Pro Val Leu Leu Ile His Gly Glu Lys Asp<br>225 230 235 240 |     |     | 720 |
| AAT GTC ATA CCT CCG GAG GCG ACC AAA AAA GCC RTAC CAA TTA ATA CCT<br>Asn Val Ile Pro Pro Glu Ala Ser Lys Lys Als Tyr Gln Leu Ile Pro<br>245 250 255    |     |     | 768 |
| TCA TTC CCT AAA GAG TTG AAA AAA TAC CCC GAT CTT GGA CAC AAC TTG<br>Ser Phe Pro Lys Glu Leu Lys Ile Tyr Pro Asp Leu Gly His Asn Leu<br>260 265 270     |     |     | 816 |
| TTT TTT GAA CCA GGC GCG GTG AAA ATC GTC ACA GAC ATT GTA GAG TGG<br>Phe Phe Glu Pro Gly Ala Val Lys Ile Val Thr Asp Ile Val Glu Trp<br>275 280 285     |     |     | 864 |
| GTT AAG AAT CTA CCC AGG GAA AAT CCT TAA<br>Val Lys Asn Leu Pro Arg Glu Asn Pro<br>290 295   |     |     | 874 |

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS
  - (A) LENGTH: 789 NUCLEOTIDES
  - (B) TYPE: NUCLEIC ACID
  - (C) STRANDEDNESS: SINGLE
  - (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

|   |     |
|---|-----|
| ATG GAG GTT TAC AAG GCC AAA TTC GGC GAA GCA AAG CTC GGC TGG GTC<br>Met Glu Val Tyr Lys Ala Lys Phe Gly Glu Ala Lys Leu Gly Trp Val<br>1 5 10 15 | 48  |
| GTT CTG GTT CAT GGC CTC GGC GAG CAC AGC GGA AGG TAT GGA AGA CTG<br>Val Leu Val His Gly Leu Gly Glu His Ser Gly Arg Tyr Gly Arg Leu<br>20 25 30  | 96  |
| ATT AAG GAA CTC AAC TAT GCC GGC TTT GGA GTT TAC ACC TTC GAC TGG<br>Ile Lys Glu Leu Asn Tyr Ala Gly Phe Gly Val Tyr Thr Phe Asp Trp              | 144 |



(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

|   |     |
|---|-----|
| TTG ATT GGC AAT TTG AAA TTG AAG AGG TTT GAA GAG GTT AAC TTA GTT<br>Leu Ile Gly Asn Leu Lys Ley Lys Arg Phe Glu Glu Val Asn Leu Val<br>1 5 10 15       | 48  |
| CTT TCG GGA GGG GCT GCC AAG GCT ATC GCC CAT ATA GGT GTT TTA AAA<br>Leu Ser Gly Gly Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys<br>20 25 30        | 96  |
| GCT CTG GAA GAG CTC GGT ATA AAG GTA AAG AGG CTC AGC GGG GTA AGT<br>Ala Leu Glu Glu Leu Gly Ile Lys Val Lys Arg Leu Ser Gly Val Ser<br>35 40 45        | 144 |
| GCT GGA GCT ATC GTT TCC GTC TTT TAC GCT TCG GGC TAC ACT CCC GAC<br>Ala Gly Ala Ile Val Ser Val Phe Tyr Ala Ser Gly Tyr Thr Pro Asp<br>50 55 60        | 192 |
| GAG ATG TTA AAA CTC CTG AAA GAG GTA AAC TGG CTC AAA CTT TTT AAG<br>Glu Met Leu Lys Leu Leu Lys Glu Val Asn Trp Leu Lys Leu Phe Lys<br>65 70 75 80     | 240 |
| TTC AAA ACA CCG AAA ATG GGC TTA ATG GGG TGG GAG AAG GCT GCA GAG<br>Phe Lys Thr Pro Lys Met Gly Leu Met Gly Trp Glu Lys Ala Ala Glu<br>85 90 95        | 288 |
| TTT TTG TAA AAA GAG CTC GGA GTT AAG AGG CTG GAA GAC CTG AAC ATA<br>Phe Leu Glu Lys Glu Leu Gly Val Lys Arg Leu Glu Asp Leu Asn Ile<br>100 105 110     | 336 |
| CCA ACC TAT CTT TGC TCG GCG GAT CTG TAC ACG GGA AAG GCT CTT TAC<br>Pro Thr Tyr Leu Cys Ser Ala Asp Ley Tyr Thr Gly Lys Ala Leu Tyr<br>115 120 125     | 384 |
| TTC GGC AGA GGT GAC TTA ATT CCC GTG CTT CTC GGA AGT TGT TCC ATA<br>Phe Gly Arg Gly Asp Leu Ile Pro Val Leu Leu Gly Ser Lys Ser Ile<br>130 135 140     | 432 |
| CCC GGG ATT TTT GAA CCA GTT GAG TAC GAG AAT TTT CTA CTT GTT GAC<br>Pro Gly Ile Phe Glu Pro Val Glu Tyr Glu Asn Phe Leu Leu Val Asp<br>145 150 155 160 | 480 |
| GGA GGT ATA GTG AAC AAC CTG CCC GTA GAA CCT TTG GAA AAG TTC AAA<br>Gly Gly Ile Val Asn Asn Leu Pro Val Glu Pro Leu Glu Lys Phe Lys<br>165 170 175     | 528 |
| GAA CCC ATA ATC GGG GTA GAT GTG CTT CCC ATA ACT CAA GAA AGA AAG<br>Glu Pro Ile Ile Gly Val Asp Val Leu Pro Ile Thr Gln Glu Arg Lys<br>180 185 190     | 576 |
| ATT AAA AAT ATA CTC CAC ATC CTT ATA AGG AGC TTC TTT CTG GCG GTT<br>Ile Lys Asn Ile Leu His Ile Leu Ile Arg Ser Phe Phe Leu Ala Val<br>195 200 205     | 624 |
| CGT TCC AAT TCG GAA AAG AGA AAG GAG TTC TGC AAC GTA GTT ATA GAA<br>Arg Ser Asn Ser Glu Lys Arg Lys Glu Phe Cys Asn Val Val Ile Glu<br>210 215 220     | 672 |
| CCT CCC CTT GAA GAG TTC TCT CCT CTG GAC GTA AAT AAG GCG GAC GAG<br>Pro Pro Leu Glu Glu Phe Ser Pro Leu Asp Val Asn Lys Ala Asp Glu                    | 720 |

225

230

235

240

ATA TTC TGC GGG GAT ATG AGA GCA CTT TAA  
 Ile Phe Cys Gly Asp Met Arg Ala Leu  
 245

730

## (2) INFORMATION FOR SEQ ID NO:30:

## (i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 1017 NUCLEOTIDES

(B) TYPE: NUCLEIC ACID

(C) STRANDEDNESS: SINGLE

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

|   |     |
|---|-----|
| ATG CCA GCT AAT GAC TCA CCC ACG ATC GAC TTT AAT CCT CGC GGC ATT     | 48  |
| Met Pro Ala Asn Asp Ser Pro Thr Ile Asp Phe Asn Pro Arg Gly Ile     |     |
| 1 5 10 15   |     |
| CTT CGC AAC GCT CAC GCA CAG GTT ATT TTA GCG ACT TCC GGC TTG CGC     | 96  |
| Leu Arg Asn Ala His Ala Gln Val Ile Leu Ala Thr Ser Gly Leu Arg     |     |
| 20 25 30  |     |
| AAA GCG TTT TTG AAA CGC ACG CAC AAG AGC TAC CTC AGC ACT GCC CAA     | 144 |
| Lys Ala Phe Leu Lys Arg Thr His Lys Ser Tyr Leu Ser Thr Ala Gln     |     |
| 35 40 45  |     |
| TGG CTG GAG CTC GAT GCC GGC AAC GGA GTT ACC TTG GCC GGA GAG CTT     | 192 |
| Trp Leu Glu Leu Asp Ala Gly Asn Gly Val Thr Leu Ala Gly Glu Leu     |     |
| 50 55 60  |     |
| AAC ACA GCG CCT GCA ACT GCA TCC TCC TCC CAC CCG GCG CAC AAG AAC 240 |     |
| Asn Thr Ala Pro Ala Thr Ala Ser Ser Ser His Pro Ala His Lys Asn     |     |
| 65 70 75 80   |     |
| ACT CTG GTT ATT GTG CTG CAC GGC TGG GAA GGC TCC AGC CAG TCG GCC     | 288 |
| Thr Leu Val Ile Val Leu His Gly Trp Glu Gly Ser Ser Gln Ser Ala     |     |
| 85 90 95  |     |
| TAT GCG ACC TCC GCT GGC AGC ACG CTT TTC GAC AAT GGG TTC GAC ACT     | 336 |
| Tyr Ala Thr Ser Ala Gly Ser Thr Leu Phe Asp Asn Gly Phe Asp Thr     |     |
| 100 105 110   |     |
| TTT CGC CTT AAT TTT CGC GAT CAC GGC GAC ACC TAC CAC TTA AAC CGC     | 384 |
| Phe Arg Leu Asn Phe Arg Asp His Gly Asp Thr Tyr His Leu Asn Arg     |     |
| 115 120 125   |     |
| GGC ATA TTT AAC TCA TCG CTG ATT GAC GAA GTA GTG GGC GCA GTC AAA     | 432 |
| Gly Ile Phe Asn Ser Ser Leu Ile Asp Glu Val Val Gly Ala Val Lys     |     |
| 130 135 140   |     |
| GCC ATC CAG CAG CAA ACC GAC TAC GAC AAG TAT TGC CTG ATG GGG TTC     | 480 |
| Ala Ile Gln Gln Gln Thr Asp Tyr Asp Lys Tyr Cys Leu Met Gly Phe     |     |
| 145 150 155 160   |     |
| TCA CTG GGT GGG AAC TTT GCC TTG CGC GTC GCG GTG CGG GAA CAG CAT     | 528 |
| Ser Leu Gly Gly Asn Phe Ala Leu Arg Val Ala Val Arg Glu Gln His     |     |
| 165 170 175   |     |

|   |       |
|---|-------|
| CTC GCT AAA CCG CTA GCG GGC GTG CTC GCC GTA TGC CCG GTA CTC GAC<br>Leu Ala Lys Pro Leu Ala Gly Val Leu Ala Val Cys Pro Val Leu Asp<br>180 185 190     | 576   |
| CCC GCA CAC ACC ATG ATG GCC CTA AAC CGA GGT GCG TTT TTC TAC GGC<br>Pro Ala His Thr Met Met Ala Leu Asn Arg Gly Ala Phe Phe Tyr Gly<br>195 200 205     | 624   |
| CGC TAT TTT GCG CAT AAA TGG AAG CGC TCG TTA ACC GCA AAA CTT GCA<br>Arg Tyr Phe Ala His Lys Trp Lys Arg Ser Leu Thr Ala Lys Leu Ala<br>210 215 220 225 | 672   |
| GCT TTC CCA GAC TAC AAA TAC GGC AAA GAT TTA AAA TCG ATA CAC ACG<br>Ala Phe Pro Asp Tyr Lys Tyr Gly Lys Asp Leu Lys Ser Ile His Thr<br>230 235 240     | 720   |
| CTT GAT GAG TTA AAC AAC TAT TTC ATT CCC CGC TAC ACC GGC TTC AAC<br>Leu Asp Glu Leu Asn Asn Tyr Phe Ile Pro Arg Tyr Thr Gly Phe Asn<br>245 250 255     | 768   |
| TCA GTC TCC GAA TAC TTC AAA AGT TAC ACG CTC ACC GGG CAG AAG CTC<br>Ser Val Ser Glu Tyr Phe Lys Ser Tyr Thr Leu Thr Gly Gln Lys Leu<br>260 265 270     | 816   |
| GCG TTT CTC AAC TGC CCC AGT TAC ATT CTG GCA GCT GGC GAC GAC CCA<br>Ala Phe Leu Asn Cys Pro Ser Tyr Ile Leu Ala Ala Gly Asp Asp Pro<br>275 280 285     | 864   |
| ATA ATT CCA GCA TCC GAC TTT CAG AAA ATA GCC AAG CCT GCG AAT CTG<br>Ile Ile Pro Ala Ser Asp Phe Gln Lys Ile Ala Lys Pro Ala Asn Leu<br>290 295 300 305 | 912   |
| CAC ATA ACA GTA ACG CAA CAA GGT TCT CAT TGC GCA TAC CTG GAA AAC<br>His Ile Thr Val Thr Gln Gln Gly Ser His Cys Ala Tyr Leu Glu Asn<br>310 315 320     | 960   |
| CTG CAT AAA CCT AGT GCT GCC GAC AAA TAT GCG GTG AAA TTA TTT GGA<br>Leu His Lys Pro Ser Ala Ala Asp Lys Tyr Ala Val Lys Leu Phe Gly<br>325 330 335     | 1,008 |
| GCC TGT TGA<br>Ala Cys  | 1,111 |

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS  
 (A) LENGTH: 936 NUCLEOTIDES  
 (B) TYPE: NUCLEIC ACID  
 (C) STRANDEDNESS: SINGLE  
 (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

|   |     |
|---|-----|
| ATG CTT GAT ATG CCA ATC GAC CCT GTT TAC TAC CAG CTT GCT GAG TAT<br>Met Leu Asp Met Pro Ile Asp Pro Val Tyr Tyr Gln Leu Ala Glu Tyr<br>1 5 10 15 | 48  |
| TTC GAC AGT CTG CCG AAG TTC GAC CAG TTT TCC TCG GCC AGA GAG TAC<br>Phe Asp Ser Leu Pro Lys Phe Asp Gln Phe Ser Ser Ala Arg Glu Tyr<br>20 25 30  | 96  |
| AGG GAG GCG ATA AAT CGA ATA TAC GAG GAG AGA AAC CGG CAG CTG AGC   | 144 |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Arg | Glu | Ala | Ile | Asn | Arg | Ile | Tyr | Glu | Glu | Arg | Asn | Arg | Gln | Leu | Ser |     |
|     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |     |
| CAG | CAT | GAG | AGG | GTT | GAA | AGA | GTT | GAG | GAC | AGG | ACG | ATT | AAG | GGG | AGG | 192 |
| Gln | His | Glu | Arg | Val | Glu | Arg | Val | Glu | Asp | Arg | Thr | Ile | Lys | Gly | Arg |     |
|     |     | 50  |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |     |
| AAC | GGA | GAC | ATC | AGA | GTC | AGA | GTT | TAC | CAG | CAG | AAG | CCC | GAT | TCC | CCG | 240 |
| Asn | Gly | Asp | Ile | Arg | Val | Arg | Val | Tyr | Gln | Gln | Lys | Pro | Asp | Ser | Pro |     |
|     |     | 65  |     |     | 70  |     |     |     |     | 75  |     |     |     |     | 80  |     |
| GGT | CTG | GTT | TAC | TAT | CAC | GGT | GGT | GGA | TTT | GTG | ATT | TGC | AGC | ATC | GAG | 288 |
| Val | Leu | Val | Tyr | Tyr | His | Gly | Gly | Gly | Phe | Val | Ile | Cys | Ser | Ile | Glu |     |
|     |     |     |     | 85  |     |     |     | 90  |     |     |     |     |     | 95  |     |     |
| TCG | CAC | GAC | GCC | TTA | TGC | AGG | AGA | AYY | GCG | AGA | CTT | TCA | AAC | TCT | ACC | 336 |
| Ser | His | Asp | Ala | Leu | Cys | Arg | Arg | Ile | Ala | Arg | Leu | Ser | Asn | Ser | Thr |     |
|     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |     |
| GTA | GTC | TCC | GTG | GAT | TAC | AGG | CTC | GCT | CCT | GAG | CAC | AAG | TTT | CCC | CCC | 384 |
| Val | Val | Ser | Val | Asp | Tyr | Arg | Leu | Ala | Pro | Glu | His | Lys | Phe | Pro | Ala |     |
|     |     | 115 |     |     |     |     | 120 |     |     |     |     | 125 |     |     |     |     |
| CCA | GTT | TAT | CAT | TGC | TAC | GAT | GCG | ACC | AAG | TGG | GTT | GCT | GAG | AAC | CGG | 432 |
| Ala | Val | Tyr | Asp | Cys | Tyr | Aso | Ala | Thr | Lys | Trp | Val | Ala | Glu | Asn | Ala |     |
|     |     | 130 |     |     |     | 135 |     |     |     |     | 140 |     |     |     |     |     |
| GAG | GAG | CTG | AGG | ATT | GAC | CCG | TCA | AAA | ATC | TTC | GTT | GGG | GGG | GAC | AGT | 480 |
| Glu | Glu | Leu | Arg | Ile | Asp | Pro | Ser | Lys | Ile | Phe | Val | Gly | Gly | Asp | Ser |     |
|     |     | 145 |     |     | 150 |     |     |     |     | 155 |     |     |     |     | 160 |     |
| GCG | GGA | CGG | AAT | CTT | GCC | CCG | GCG | CTT | TCA | ATA | ATG | GCG | AGA | GAC | AGC | 528 |
| Ala | Gly | Gly | Asn | Leu | Ala | Ala | Ala | Val | Ser | Ile | Met | Ala | Arg | Asp | Ser |     |
|     |     |     |     | 165 |     |     |     |     | 170 |     |     |     |     | 175 |     |     |
| GGA | GAA | GAT | TTC | ATA | AAG | CAT | CAA | ATT | CTA | ACT | TAC | CCC | GTT | GTG | AAC | 576 |
| Gly | Glu | Asp | Phe | Ile | Lys | His | Gln | Ile | Leu | Ile | Tyr | Pro | Val | Val | Asn |     |
|     |     |     | 180 |     |     |     |     | 185 |     |     |     |     | 190 |     |     |     |
| TTT | GTA | GCC | CCC | ACA | CCA | TCG | CTT | CTG | GAG | TTT | GGA | GAG | GGG | CTG | TGG | 624 |
| Phe | Val | Ala | Pro | Thr | Pro | Ser | Leu | Leu | Glu | Phe | Gly | Glu | Gly | Leu | Trp |     |
|     |     | 195 |     |     |     |     | 200 |     |     |     |     | 205 |     |     |     |     |
| ATT | CTC | GAC | CAG | AAG | ATA | ATG | AGT | TGG | TTC | TCG | GAG | CAG | TAC | TTC | TCC | 672 |
| Ile | Leu | Asp | Gln | Lys | Ile | Met | Ser | Trp | Phe | Ser | Glu | Gln | Tyr | Phe | Ser |     |
|     |     | 210 |     |     |     | 215 |     |     |     |     | 230 |     |     |     |     |     |
| AGA | GAG | GAA | GAT | AAG | TTC | AAG | CCC | CTC | GCC | TCC | GTA | ATC | TTT | GCG | GAC | 720 |
| Arg | Glu | Glu | Aso | Lys | Phe | Asn | Pro | Leu | Ala | Ser | Val | Ile | Phe | Ala | Asp |     |
|     |     | 235 |     |     | 240 |     |     |     |     | 245 |     |     |     |     | 250 |     |
| CTT | GAG | AAC | CTA | CCT | CCT | GCG | CTG | ATC | ATA | ACC | GCC | GAA | TAC | GAC | CCG | 768 |
| Leu | Glu | Asn | Leu | Pro | Pro | Ala | Leu | Ile | Ile | Thr | Ala | Glu | Tyr | Asp | Pro |     |
|     |     |     |     | 255 |     |     |     |     | 260 |     |     |     |     | 265 |     |     |
| CTG | AGA | GAT | GAA | GGA | GAA | GTT | TTC | GGG | CAG | ATG | CTG | AGA | AGA | GCC | GGT | 816 |
| Leu | Arg | Asp | Glu | Gly | Glu | Val | Phe | Gly | Gln | Met | Leu | Arg | Arg | Ala | Gly |     |
|     |     |     | 270 |     |     |     |     | 275 |     |     |     |     | 280 |     |     |     |
| GTT | GAG | GCG | AGC | ATC | GTC | AGA | TAC | AGA | GGC | GTG | CTT | CAC | GGA | TTC | ATC | 864 |
| Val | Glu | Ala | Ser | Ile | Val | Arg | Tyr | Arg | Gly | Val | Leu | His | Gly | Phe | Ile |     |
|     |     | 285 |     |     |     |     | 290 |     |     |     |     | 295 |     |     |     |     |
| AAT | TAC | TAT | CCC | GTG | CTG | AAG | GCT | GCG | AGG | GAT | GCG | ATA | AAC | CAG | ATT | 912 |
| Asn | Tyr | Tyr | Pro | Val | Leu | Lys | Ala | Ala | Arg | Asp | Ala | Ile | Asn | Gln | Ile |     |



|                                 |  |     |  |     |     |
|---------------------------------|--|-----|--|-----|-----|
| 300                             |  | 305 |  | 310 |     |
| GCC GCT CTT CTT GTG TTC GAC TAG |  |     |  |     | 936 |
| Ala Ala Leu leu Val Phe Asp     |  |     |  |     |     |
| 315                             |  | 320 |  |     |     |

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS  
 (A) LENGTH: 918 NUCLEOTIDES  
 (B) TYPE: NUCLEIC ACID  
 (C) STRANDEDNESS: SINGLE  
 (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

|   |     |
|---|-----|
| ATG CCC CTA GAT CCT AGA ATT AAA AAG TTA CTA GAA TCA GCT CTT ACT | 48  |
| Met Pro Leu Asp Pro Arg Ile Lys Lys Leu Leu Glu Ser Ala Leu Thr |     |
| 5 10 15   |     |
| ATA CCA ATT GGT AAA GCC CCA GTA GAA GAG GTA AGA AAG ATA TTT AGG | 96  |
| Ile Pro Ile Gly Lys Ala Pro Val Glu Glu Val Arg Lys Ile Phe Arg |     |
| 20 25 30  |     |
| CAA TTA GCG TCG GCA GCT CCC AAA GTC GAA GTT GGA AAA GTA GAA GAT | 144 |
| Gln Leu Ala Ser Ala Ala Pro Lys Val Glu Val Gly Lys Val Glu Asp |     |
| 35 40 45  |     |
| ATA AAA ATA CCA GGC AGT GAA ACC GTT ATA AAC GCT AGA GTG TAT TTT | 192 |
| Ile Lys Ile Pro Gly Ser Glu Thr Val Ile Asn Ala Arg Val Tyr Phe |     |
| 50 55 60  |     |
| CCG AAG AGT AGC GGT CCT TAT GGT GTT CTA GTG TAT CTT CAT GGA GGC | 240 |
| Pro Lys Ser Ser Gly Pro Tyr Gly Val Leu Val Tyr Leu His Gly Gly |     |
| 65 70 75 80   |     |
| GGT TTT GTA ATA GGC GAT GTG GAA TCT TAT GAC CCA TTA TGT AGA GCA | 288 |
| Gly Phe Val Ile Gly Asp Val Glu Ser Tyr Asp Pro Leu Cys Arg Ala |     |
| 85 90 95  |     |
| ATT ACA AAT GCG TGC AAT TGC GTT GTA GTA TCA GTG GAC TAT AGG TTA | 336 |
| Ile Thr Asn Ala Cys Asn Cys Val Val Ser Val Asp Tyr Arg Leu     |     |
| 100 105 110   |     |
| GCT CCA GAA TAC AAG TTT CCT TCT GCA GTT ATC GAT TCA TTT GAC GCT | 384 |
| Ala Pro Glu Tyr Lys Phe Pro Ser Ala Val Ile Asp Ser Phe Asp Ala |     |
| 115 120 125   |     |
| ACT AAT TGG GTT TAT AAC AAT TTA GAT AAA TTT GAT GGA AAG ATG GGA | 432 |
| Thr Asn Trp Val Tyr Asn Asn Leu Asp Lys Phe Asp Gly Lys Met Gly |     |
| 130 135 140   |     |
| GTT GCG ATT GCG GGA GAT AGT GCT GGA GGA AAT TTG GCA GCG GTT GTA | 480 |
| Val Ala Ile Ala Gly Asp Ser Ale Gly Gly Asn Leu Ala Ala Val Val |     |
| 145 150 155 160   |     |
| GCT CTT CTT TCA AAG GGT AAA ATT AAT TTG AAG TAT CAA ATA CTG GTT | 528 |
| Ala Leu Leu Ser Lys Gly Lys Ile Asn Leu Lys Tyr Gln Ile Leu Val |     |
| 165 170 175   |     |
| TAC CCA GCG GTA AGT TTA GAT AAC GTT TCA AGA TCC ATG ATA GAG TAC | 576 |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|-----|
| Tyr | Pro | Ala | Val | Ser | Leu | Asp | Asn | Val | Ser | Arg | Ser | Met | Ile | Glu | Tyr |  |     |
|     |     |     | 180 |     |     |     |     | 185 |     |     |     |     | 190 |     |     |  |     |
| TCT | GAT | GGG | TTC | TTC | CTT | ACC | ACA | GAG | CAT | ATA | GAG | TGG | TTC | GGT | TCT |  | 624 |
| Ser | Asp | Gly | Phe | Phe | Leu | Thr | Arg | Glu | His | Ile | Glu | Trp | Phe | Gly | Ser |  |     |
|     |     | 195 |     |     |     |     | 200 |     |     |     | 205 |     |     |     |     |  |     |
| CAA | TAC | TTA | CGA | AGC | CCT | GCA | GAT | TTG | CTA | GAC | TTT | AGG | TTC | TCT | CCA |  | 672 |
| Gln | Tyr | Leu | Arg | Ser | Pro | Ala | Asp | Leu | Leu | Asp | Phe | Arg | Phe | Ser | Pro |  |     |
|     | 210 |     |     |     |     | 215 |     |     |     |     | 220 |     |     |     |     |  |     |
| ATT | CTG | GCG | CAA | GAT | TTC | AAC | GCA | TTA | CCT | CCA | GCC | TTG | ATA | ATA | ACA |  | 720 |
| Ile | Leu | Ala | Gln | Asp | Phe | Asn | Gly | Leu | Pro | Pro | Ala | Leu | Ile | Ile | Thr |  |     |
|     | 225 |     |     |     | 230 |     |     |     | 235 |     |     |     |     |     | 240 |  |     |
| GCA | GAA | TAC | GAT | CCA | CTA | AGG | GAT | CAA | GGA | GAA | GCG | TAT | GCA | AAT | AAA |  | 768 |
| Ala | Glu | Tyr | Asp | Pro | Leu | Arg | Asp | Gln | Gly | Glu | Ala | Tyr | Ala | Asn | Lys |  |     |
|     |     |     |     | 245 |     |     |     | 250 |     |     |     |     |     | 255 |     |  |     |
| CTA | CTA | CAA | GCT | GGA | GTC | TCA | GTT | ACT | AGT | GTG | AGA | TTT | AAC | AAC | GTT |  | 816 |
| Leu | Leu | Gln | Ala | Gly | Val | Ser | Val | Thr | Ser | Val | Arg | Phe | Asn | Asn | Val |  |     |
|     |     | 260 |     |     |     |     | 265 |     |     |     |     |     | 270 |     |     |  |     |
| ATA | CAC | GGA | TTC | CTC | TCA | TTC | TTT | CCG | TTG | ATG | GAG | CAA | GGA | AGA | GAT |  | 864 |
| Ile | His | Gly | Phe | Leu | Ser | Phe | Phe | Pro | Leu | Met | Glu | Gln | Gly | Arg | Asp |  |     |
|     |     | 275 |     |     |     |     | 280 |     |     |     |     | 285 |     |     |     |  |     |
| GCT | ATA | GGT | CTG | ATA | GGG | TCT | GTG | TTA | AGA | CGA | GTA | TTT | TAT | GAT | AAA |  | 912 |
| Ala | Ile | Gly | Leu | Ile | Gly | Ser | Val | Leu | Arg | Arg | Val | Phe | Tyr | Asp | Lys |  |     |
|     | 290 |     |     |     |     | 295 |     |     |     |     | 300 |     |     |     |     |  |     |
| ATT | TAA |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  | 918 |
| Ile |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |     |
| 305 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |     |

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 184 AMINO ACIDS

(B) TYPE: AMINO ACID

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|
| Met | Ser | Leu | Asn | Lys | His | Ser | Trp | Met | Asp | Met | Ile | Ile | Phe | Ile | Leu |  |  |
| 1   |     |     |     | 5   |     |     |     | 10  |     |     |     |     | 15  |     |     |  |  |
| Ser | Phe | Ser | Phe | Pro | Leu | Thr | Met | Ile | Ala | Leu | Ala | Ile | Ser | Met | Ser |  |  |
|     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |     |  |  |
| Ser | Trp | Phe | Asn | Ile | Trp | Asn | Asn | Ala | Leu | Ser | Asp | Leu | Gly | His | Ala |  |  |
|     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |     |  |  |
| Val | Lys | Ser | Ser | Val | Ala | Pro | Ile | Phe | Asn | Leu | Gly | Leu | Ala | Ile | Gly |  |  |
|     | 50  |     |     |     | 55  |     |     |     |     |     | 60  |     |     |     |     |  |  |
| Gly | Ile | Leu | Ile | Val | Ile | Val | Gly | Leu | Arg | Asn | Leu | Tyr | Ser | Trp | Ser |  |  |
|     | 65  |     |     | 70  |     |     |     |     | 75  |     |     |     |     | 80  |     |  |  |
| Arg | Val | Lys | Gly | Ser | Leu | Ile | Ile | Ser | Met | Gly | Val | Phe | Leu | Asn | Leu |  |  |
|     |     |     | 85  |     |     |     |     | 90  |     |     |     |     |     | 95  |     |  |  |

Ile Gly Val Phe Asp Glu Val Tyr Gly Trp Ile His Phe Leu Val Ser  
100 105 110  
Val Leu Phe Phe Leu Ser Ile Ile Ala Tyr Phe Ile Ala Ile Ser Ile  
115 120 125  
Leu Asp Lys Ser Trp Ile Ala Val Leu Leu Ile Ile Gly His Ile Ala  
130 135 140  
Met Trp Tyr Leu His Phe Ala Ser Glu Ile Pro Arg Gly Ala Ala Ile  
145 150 155 160  
Pro Glu Leu Leu Ala Val Phe Ser Phe Leu Pro Phe Tyr Ile Arg Asp  
165 170 175  
Tyr Phe Lys Ser Tyr Thr Lys Arg  
180

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 346 AMINO ACIDS

(B) TYPE: AMINO ACID

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Met Lys Leu Leu Glu Pro Thr Asn Thr Ser Tyr Thr Leu Leu Gln Asp  
1 5 10 15  
Leu Ala Leu His Phe Ala Phe Tyr Trp Phe Leu Ala Val Tyr Thr Trp  
20 25 30  
Leu Pro Gly Val Leu Val Arg Gly Val Ala Val Asp Thr Gly Val Ala  
35 40 45  
Arg Val Pro Gly Leu Gly Arg Arg Gly Lys Arg Leu Leu Leu Ala Ala  
50 55 60  
Val Ala Val Leu Ala Leu Val Val Ser Val Val Val Pro Ala Tyr Val  
65 70 75 80  
Ala Tyr Ser Ser Leu His Pro Glu Ser Cys Arg Pro Val Ala Pro Glu  
85 90 95  
Gly Leu Thr Tyr Lys Glu Phe Ser Val Thr Ala Glu Asp Gly Leu Val  
100 105 110  
Val Arg Gly Trp Cal Leu Gly Pro Gly Ala Gly Gly Asn Pro Val Phe  
115 120 125  
Val Leu Met His Gly Tyr Thr Gly Cys Arg Ser Ala Pro Tyr Met Ala  
130 135 140  
Val Leu Ala Arg Glu Leu Val Glu Trp Gly Tyr Pro Val Val Val Phe  
145 150 155 160  
Asp Phe Arg Gly His Gly Glu Ser Gly Gly Ser Thr Thr Ile Gly Pro  
165 170 175  
Arg Glu Val Leu Asp Ala Arg Ala Val Val Gly Tyr Val Ser Glu Arg  
180 185 190

Phe Pro Gly Arg Arg Ile Ile Leu Val Gly Phe Ser Met Gly Gly Ala  
 195 200 205  
 Val Ala Ile Val Glu Gly Ala Gly Asp Pro Arg Val Tyr Ala Val Ala  
 210 215 220  
 Ala Asp Ser Pro Tyr Tyr Arg Leu Arg Asp Val Ile Pro Arg Trp Leu  
 225 230 235 240  
 Glu Tyr Lys Thr Pro Leu Pro Gly Trp Val Gly Val Leu Ala Gly Phe  
 245 250 255  
 Tyr Gly Arg Leu Met Ala Gly Val Asp Leu Gly Phe Gly Pro Ala Gly  
 260 265 270  
 Val Gly Arg Val Asp Lys Pro Leu Leu Val Val Tyr Gly Pro Arg Asp  
 275 280 285  
 Pro Leu Val Thr Arg Asp Glu Ala Arg Ser Leu Ala Ser Arg Ser Pro  
 290 295 300  
 Cys Gly Arg Leu Val Glu Val Pro Gly Ala Gly His Val Glu Ala Val  
 305 310 315 320  
 Asp Val Leu Gly Pro Gly Arg Tyr Ala Asp Met Leu Ile Glu Leu Ala  
 325 330 335  
 His Glu Glu Cys Pro Pro Gly Ala Gly Gly  
 340 345

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 262 AMINO ACIDS

(B) TYPE: AMINO ACID

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Met Pro Tyr Val Arg Asn Gly Gly Val Asn Ile Tyr Tyr Glu Leu Val  
 1 5 10 15  
 Asp Gly Pro Glu Pro Pro Ile Val Phe Val His Gly Trp Thr Ala Asn  
 20 25 30  
 Met Asn Phe Trp Lys Glu Gln Arg Arg Tyr Phe Ala Gly Arg Asn Met  
 35 40 45  
 Met Leu Phe Val Asp Asn Arg Gly His Gly Arg Ser Asp Lys Pro Leu  
 50 55 60  
 Gly Tyr Asp Phe Tyr Arg Phe Glu Asn Phe Ile Ser Asp Leu Asp Ala  
 65 70 75 80  
 Val Val Arg Glu Thr Gly Val Glu Lys Phe Cal Leu Val Gly His Ser  
 85 90 95  
 Phe Gly Thr Met Ile Ser Met Lys Tyr Cys Ser Glu Tyr Arg Asn Arg  
 100 105 110  
 Val Leu Ala Leu Ile Leu Ile Gly Gly Gly Ser Arg Ile Lys Leu Leu

|                         |                     |                     |
|-------------------------|---------------------|---------------------|
| 115                     | 120                 | 125                 |
| His Arg Ile Gly Tyr Pro | Leu Ala Lys Ile Leu | Ala Ser Ile Ala Tyr |
| 130                     | 135                 | 140                 |
| Lys Lys Ser Ser Arg Leu | Val Ala Asp Leu Ser | Phe Gly Lys Asn Ala |
| 145                     | 150                 | 155                 |
| Gly Glu Leu Lys Glu Trp | Gly Trp Lys Gln Ala | Met Asp Tyr Thr Pro |
| 165                     | 170                 | 175                 |
| Ser Tyr Val Ala Met Tyr | Thr Tyr Arg Thr Leu | Thr Lys Val Asn Leu |
| 180                     | 185                 | 190                 |
| Glu Asn Ile Leu Glu Lys | Ile Asp Cys Pro Thr | Leu Ile Ile Val Gly |
| 195                     | 200                 | 205                 |
| Glu Glu Asp Ala Leu Leu | Pro Val Ser Lys Ser | Val Glu Leu Ser Arg |
| 210                     | 215                 | 220                 |
| Arg Ile Glu Asn Ser Lys | Leu Val Ile Ile Pro | Asn Ser Gly His Cys |
| 225                     | 230                 | 235                 |
| Val Met Leu Glu Ser Pro | Ser Glu Val Asn Arg | Ala Met Asp Glu Phe |
| 245                     | 250                 | 255                 |
| Ile Ser Ser Ala Gln Phe |                     |                     |
| 260                     |                     |                     |

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 251 AMINO ACIDS
- (B) TYPE: AMINO ACID
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

|   |
|---|
| Leu Arg Leu Arg Lys Phe Glu Glu Ile Asn Leu Val Leu Ser Gly Gly |
| 1 5 10 15   |
| Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys Ala Ile Asn Glu |
| 20 25 30  |
| Leu Glu Ile Arg Val Arg Ala Leu Ser Gly Val Ser Ala Gly Ala Ile |
| 35 40 45  |
| Val Ser Val Phe Tyr Ala Ser Gly Tyr Ser Pro Glu Gly Met Phe Ser |
| 50 55 60  |
| Leu Leu Lys Arg Val Asn Trp Leu Lys Leu Phe Lys Phe Lys Pro Pro |
| 65 70 75 80   |
| Leu Lys Gly Leu Ile Gly Trp Glu Lys Ala Ile Arg Phe Leu Glu Glu |
| 85 90 95  |
| Val Leu Pro Tyr Arg Arg Ile Glu Lys Leu Glu Ile Pro Thr Tyr Ile |
| 100 105 110   |
| Cys Ala Thr Asp Leu Tyr Ser Gly Arg Ala Leu Tyr Leu Ser Glu Gly |
| 115 120 125   |

Ser Leu Ile Pro Ala Leu Leu Gly Ser Cys Ala Ile Pro Gly Ile Phe  
130 135 140

Glu Pro Val Glu Tyr Lys Asn Tyr Leu Leu Val Asp Gly Gly Ile Val  
145 150 155 160

Asn Asn Leu Pro Val Glu Pro Phe Gln Glu Ser Gly Ile Pro Thr Val  
165 170 175

Cys Val Asp Val Leu Pro Ile Glu Pro Glu Lys Asp Ile Lys Asn Ile  
180 185 190

Leu His Ile Leu Leu Arg Ser Phe Phe Leu Ala Val Arg Ser Asn Ser  
195 200 205

Glu Lys Arg Lys Glu Phe Cys Asp Leu Val Ile Val Pro Glu Leu Glu  
210 215 220

Glu Phe Thr Pro Leu Asp Val Arg Lys Ala Asp Gln Ile Met Glu Arg  
225 230 235 240

Gly Tyr Ile Lys Ala Leu Glu Val Leu Ser Glu  
245 250

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 297 AMINO ACIDS
- (B) TYPE: AMINO ACID
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Met Phe Asn Ile Asn Val Phe Val Asn Ile Ser Trp Leu Tyr Phe Ser  
1 5 10 15

Gly Ile Val Met Lys Thr Val Glu Glu Tyr Ala Leu Leu Glu Thr Gly  
20 25 30

Val Arg Val Phe Tyr Arg Cys Val Ile Pro Glu Lys Ala Phe Asn Thr  
35 40 45

Leu Ile Ile Gly Ser His Gly Leu Gly Ala His Ser Gly Ile Tyr Ile  
50 55 60

Ser Val Ala Glu Glu Phe Ala Arg His Gly Phe Gly Phe Cys Met His  
65 70 75 80

Asp Gln Arg Gly His Gly Arg Thr Ala Ser Asp Arg Glu Arg Gly Tyr  
85 90 95

Val Glu Gly Phe His Asn Phe Ile Glu Asp Met Lys Ala Phe Ser Asp  
100 105 110

Tyr Ala Lys Trp Arg Val Gly Gly Asp Glu Ile Ile Leu Leu Gly His  
115 120 125

Ser Met Gly Gly Leu Ile Ala Leu Leu Thr Val Ala Thr Tyr Lys Glu  
130 135 140

Ile Ala Lys Gly Val Ile Ala Leu Ala Pro Ala Leu Gln Ile Pro Leu  
145 150 155 160

Thr Pro Ala Arg Arg Leu Val Leu Ser Leu Ala Ser Arg Leu Ala Pro  
 165 170 175  
 His Ser Lys Ile Thr Leu Gln Arg Arg Leu Pro Gln Lys Pro Glu Gly  
 180 185 190  
 Phe Gln Arg Ala Lys Asp Ile Glu Tyr Ser Leu Ser Glu Ile Ser Val  
 195 200 205  
 Lys Leu Val Asp Glu Met Ile Lys Ala Ser Ser Met Phe Trp Thr Ile  
 210 215 220  
 Ala Gly Glu Ile Asn Thr Pro Val Leu Leu Ile His Gly Glu Lys Asp  
 225 230 235 240  
 Asn Val Ile Pro Pro Glu Ala Ser Lys Lys Als Tyr Gln Leu Ile Pro  
 245 250 255  
 Ser Phe Pro Lys Glu Leu Lys Ile Tyr Pro Asp Leu Gly His Asn Leu  
 260 265 270  
 Phe Phe Glu Pro Gly Ala Val Lys Ile Val Thr Asp Ile Val Glu Trp  
 275 280 285  
 Val Lys Asn Leu Pro Arg Glu Asn Pro  
 290 295

(2) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS  
 (A) LENGTH: 262 AMINO ACIDS  
 (B) TYPE: AMINO ACID  
 (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Met Glu Val Tyr Lys Ala Lys Phe Gly Glu Ala Lys Leu Gly Trp Val  
 1 5 10 15  
 Val Leu Val His Gly Leu Gly Glu His Ser Gly Arg Tyr Gly Arg Leu  
 20 25 30  
 Ile Lys Glu Leu Asn Tyr Ala Gly Phe Gly Val Tyr Thr Phe Asp Trp  
 35 40 45  
 Pro Gly His Gly Lys Ser Pro Gly Lys Arg Gly His Thr Ser Val Glu  
 50 55 60  
 Glu Ala Met Glu Ile Ile Asp Ser Ile Ile Glu Glu Ile Arg Glu Lys  
 65 70 75 80  
 Pro Phe Leu Phe Gly His Ser Leu Gly Gly Leu Thr Val Ile Arg Tyr  
 85 90 95  
 Ala Glu Thr Arg Pro Asp Lys Ile Arg Gly Leu Ile Ala Ser Ser Pro  
 100 105 110  
 Ala Leu Ala Lys Ser Pro Glu Thr Pro Gly Phe Met Val Ala Leu Ala  
 115 120 125  
 Lys Phe Leu Gly Lys Ile Ala Pro Gly Val Val Leu Ser Asn Gly Ile

130                      135                      140  
 Lys Pro Glu Leu Leu Ser Arg Asn Arg Asp Ala Val Arg Arg Tyr Val  
 145                      150                      155                      160  
 Glu Asp Pro Leu Val His Asp Arg Ile Ser Ala Lys Leu Gly Arg Ser  
                     165                      170                      175  
 Ile Phe Val Asn Met Glu Leu Ala His Arg Glu Ala Asp Lys Ile Lys  
                     180                      185                      190  
 Val Pro Ile Leu Leu Leu Ile Gly Thr Gly Asp Val Ile Thr Pro Pro  
                     195                      200                      205  
 Glu Gly Ser Arg Arg Leu Phe Glu Glu Leu Ala Val Glu Asn Lys Thr  
                     210                      215                      220  
 Leu Arg Glu Phe Glu Gly Ala Tyr His Glu Ile Phe Glu Asp Pro Glu  
 225                      230                      235                      240  
 Trp Ala Glu Glu Phe His Glu Thr Ile Val Lys Trp Leu Val Glu Lys  
                     245                      250                      255  
 Ser Tyr Ser Ser Ala Gln  
                     260

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 249 AMINO ACIDS

(B) TYPE: AMINO ACID

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Leu Ile Gly Asn Leu Lys Ley Lys Arg Phe Glu Glu Val Asn Leu Val  
 1                      5                      10                      15  
 Leu Ser Gly Gly Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys  
                     20                      25                      30  
 Ala Leu Glu Glu Leu Gly Ile Lys Val Lys Arg Leu Ser Gly Val Ser  
                     35                      40                      45  
 Ala Gly Ala Ile Val Ser Val Phe Tyr Ala Ser Gly Tyr Thr Pro Asp  
                     50                      55                      60  
 Glu Met Leu Lys Leu Leu Lys Glu Val Asn Trp Leu Lys Leu Phe Lys  
 65                      70                      75                      80  
 Phe Lys Thr Pro Lys Met Gly Leu Met Gly Trp Glu Lys Ala Ala Glu  
                     85                      90                      95  
 Phe Leu Glu Lys Glu Leu Gly Val Lys Arg Leu Glu Asp Leu Asn Ile  
                     100                      105                      110  
 Pro Thr Tyr Leu Cys Ser Ala Asp Ley Tyr Thr Gly Lys Ala Leu Tyr  
                     115                      120                      125  
 Phe Gly Arg Gly Asp Leu Ile Pro Val Leu Leu Gly Ser Lys Ser Ile  
 130                      135                      140



Pro Gly Ile Phe Glu Pro Val Glu Tyr Glu Asn Phe Leu Leu Val Asp  
 145 150 155 160  
 Gly Gly Ile Val Asn Asn Leu Pro Val Glu Pro Leu Glu Lys Phe Lys  
 165 170 175  
 Glu Pro Ile Ile Gly Val Asp Val Leu Pro Ile Thr Gln Glu Arg Lys  
 180 185 190  
 Ile Lys Asn Ile Leu His Ile Leu Ile Arg Ser Phe Phe Leu Ala Val  
 195 200 205  
 Arg Ser Asn Ser Glu Lys Arg Lys Glu Phe Cys Asn Val Val Ile Glu  
 210 215 220  
 Pro Pro Leu Glu Glu Phe Ser Pro Leu Asp Val Asn Lys Ala Asp Glu  
 225 230 235 240  
 Ile Phe Cys Gly Asp Met Arg Ala Leu  
 245

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 339 AMINO ACIDS

(B) TYPE: AMINO ACID

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Met Pro Ala Asn Asp Ser Pro Thr Ile Asp Phe Asn Pro Arg Gly Ile  
 1 5 10 15  
 Leu Arg Asn Ala His Ala Gln Val Ile Leu Ala Thr Ser Gly Leu Arg  
 20 25 30  
 Lys Ala Phe Leu Lys Arg Thr His Lys Ser Tyr Leu Ser Thr Ala Gln  
 35 40 45  
 Trp Leu Glu Leu Asp Ala Gly Asn Gly Val Thr Leu Ala Gly Glu Leu  
 50 55 60  
 Asn Thr Ala Pro Ala Thr Ala Ser Ser Ser His Pro Ala His Lys Asn  
 65 70 75 80  
 Thr Leu Val Ile Val Leu His Gly Trp Glu Gly Ser Ser Gln Ser Ala  
 85 90 95  
 Tyr Ala Thr Ser Ala Gly Ser Thr Leu Phe Asp Asn Gly Phe Asp Thr  
 100 105 110  
 Phe Arg Leu Asn Phe Arg Asp His Gly Asp Thr Tyr His Leu Asn Arg  
 115 120 125  
 Gly Ile Phe Asn Ser Ser Leu Ile Asp Glu Val Val Gly Ala Val Lys  
 130 135 140  
 Ala Ile Gln Gln Gln Thr Asp Tyr Asp Lys Tyr Cys Leu Met Gly Phe  
 145 150 155 160  
 Ser Leu Gly Gly Asn Phe Ala Leu Arg Val Ala Val Arg Glu Gln His  
 165 170 175

Leu Ala Lys Pro Leu Ala Gly Val Leu Ala Val Cys Pro Val Leu Asp  
 180 185 190  
 Pro Ala His Thr Met Met Ala Leu Asn Arg Gly Ala Phe Phe Tyr Gly  
 195 200 205  
 Arg Tyr Phe Ala His Lys Trp Lys Arg Ser Leu Thr Ala Lys Leu Ala  
 210 215 220 225  
 Ala Phe Pro Asp Tyr Lys Tyr Gly Lys Asp Leu Lys Ser Ile His Thr  
 230 235 240  
 Leu Asp Glu Leu Asn Asn Tyr Phe Ile Pro Arg Tyr Thr Gly Phe Asn  
 245 250 255  
 Ser Val Ser Glu Tyr Phe Lys Ser Tyr Thr Leu Thr Gly Gln Lys Leu  
 260 265 270  
 Ala Phe Leu Asn Cys Pro Ser Tyr Ile Leu Ala Ala Gly Asp Asp Pro  
 275 280 285  
 Ile Ile Pro Ala Ser Asp Phe Gln Lys Ile Ala Lys Pro Ala Asn Leu  
 290 295 300 305  
 His Ile Thr Val Thr Gln Gln Gly Ser His Cys Ala Tyr Leu Glu Asn  
 310 315 320  
 Leu His Lys Pro Ser Ala Ala Asp Lys Tyr Ala Val Lys Leu Phe Gly  
 325 330 335  
 Ala Cys

(2) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS  
 (A) LENGTH: 311 AMINO ACIDS  
 (B) TYPE: AMINO ACID  
 (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Met Leu Asp Met Pro Ile Asp Pro Val Tyr Tyr Gln Leu Ala Glu Tyr  
 1 5 10 15  
 Phe Asp Ser Leu Pro Lys Phe Asp Gln Phe Ser Ser Ala Arg Glu Tyr  
 20 25 30  
 Arg Glu Ala Ile Asn Arg Ile Tyr Glu Glu Arg Asn Arg Gln Leu Ser  
 35 40 45  
 Gln His Glu Arg Val Glu Arg Val Glu Asp Arg Thr Ile Lys Gly Arg  
 50 55 60  
 Asn Gly Asp Ile Arg Val Arg Val Tyr Gln Gln Lys Pro Asp Ser Pro  
 65 70 75 80  
 Val Leu Val Tyr Tyr His Gly Gly Gly Phe Val Ile Cys Ser Ile Glu  
 85 90 95  
 Ser His Asp Ala Leu Cys Arg Arg Ile Ala Arg Leu Ser Asn Ser Thr  
 100 105 110

Val Val Ser Val Asp Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Ala  
 115 120 125  
 Ala Val Tyr Asp Cys Tyr Aso Ala Thr Lys Trp Val Ala Glu Asn Ala  
 130 135 140  
 Glu Glu Leu Arg Ile Asp Pro Ser Lys Ile Phe Val Gly Gly Asp Ser  
 145 150 155 160  
 Ala Gly Gly Asn Leu Ala Ala Ala Val Ser Ile Met Ala Arg Asp Ser  
 165 170 175  
 Gly Glu Asp Phe Ile Lys His Gln Ile Leu Ile Tyr Pro Val Val Asn  
 180 185 190  
 Phe Val Ala Pro Thr Pro Ser Leu Leu Glu Phe GLy Glu Gly Leu Trp  
 195 200 205  
 Ile Leu Asp Gln Lys Ile Met Ser Trp Phe Ser Glu Gln Tyr Phe Ser  
 210 215 230  
 Arg Glu Glu Aso Lys Phe Asn Pro Leu Ala Ser Val Ile Phe Ala Asp  
 235 240 245 250  
 Leu Glu Asn Leu Pro Pro Ala Leu Ile Ile Thr Ala Glu Tyr Asp Pro  
 255 260 265  
 Leu Arg Asp Glu Gly Glu Val Phe Gly Gln Met Leu Arg Arg Ala Gly  
 270 275 280  
 Val Glu Ala Ser Ile Val Arg Tyr Arg Gly Val Leu His Gly Phe Ile  
 285 290 295  
 Asn Tyr Tyr Pro Val Leu Lys Ala Ala Arg Asp Ala Ile Asn Gln Ile  
 300 305 310  
 Ala Ala Leu leu Val Phe Asp  
 315 320

(2) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS
  - (A) LENGTH: 305 AMINO ACIDS
  - (B) TYPE: AMINO ACID
  - (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Met Pro Leu Asp Pro Arg Ile Lys Lys Leu Leu Glu Ser Ala Leu Thr  
 5 10 15  
 Ile Pro Ile Gly Lys Ala Pro Val Glu Glu Val Arg Lys Ile Phe Arg  
 20 25 30  
 Gln Leu Ala Ser Ala Ala Pro Lys Val Glu Val Gly Lys Val Glu Asp  
 35 40 45  
 Ile Lys Ile Pro Gly Ser Glu Thr Val Ile Asn Ala Arg Val Tyr Phe  
 50 55 60

Pro Lys Ser Ser Gly Pro Tyr Glu Val Leu Val Tyr Leu His Gly Gly  
 65 70 75 80  
 Gly Phe Val Ile Gly Asp Val Glu Ser Tyr Asp Pro Leu Cys Arg Ala  
 85 90 95  
 Ile Thr Asn Ala Cys Asn Cys Val Val Ser Val Asp Tyr Arg Leu  
 100 105 110  
 Ala Pro Glu Tyr Lys Phe Pro Ser Ala Val Ile Asp Ser Phe Asp Ala  
 115 120 125  
 Thr Asn Trp Val Tyr Asn Asn Leu Asp Lys Phe Asp Gly Lys Met Gly  
 130 135 140  
 Val Ala Ile Ala Gly Asp Ser Ala Gly Gly Asn Leu Ala Ala Val Val  
 145 150 155 160  
 Ala Leu Leu Ser Lys Gly Lys Ile Asn Leu Lys Tyr Gln Ile Leu Val  
 165 170 175  
 Tyr Pro Ala Val Ser Leu Asp Asn Val Ser Arg Ser Met Ile Glu Tyr  
 180 185 190  
 Ser Asp Gly Phe Phe Leu Thr Arg Glu His Ile Glu Trp Phe Gly Ser  
 195 200 205  
 Gln Tyr Leu Arg Ser Pro Ala Asp Leu Leu Asp Phe Arg Phe Ser Pro  
 210 215 220  
 Ile Leu Ala Gln Asp Phe Asn Gly Leu Pro Pro Ala Leu Ile Ile Thr  
 225 230 235 240  
 Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Ala Tyr Ala Asn Lys  
 245 250 255  
 Leu Leu Gln Ala Gly Val Ser Val Thr Ser Val Arg Phe Asn Asn Val  
 260 265 270  
 Ile His Gly Phe Leu Ser Phe Phe Pro Leu Met Glu Gln Gly Arg Asp  
 275 280 285  
 Ala Ile Gly Leu Ile Gly Ser Val Leu Arg Arg Val Phe Tyr Asp Lys  
 290 295 300  
 Ile  
 305

**What Is Claimed Is:**

1. An isolated polynucleotide comprising a member selected from the group consisting of:

(a) a polynucleotide having at least a 70% identity to a polynucleotide encoding an enzyme comprising amino acid sequences set forth in SEQ ID NOS:33-42;

(b) a polynucleotide which is complementary to the polynucleotide of (a);  
and

(c) a polynucleotide comprising at least 15 consecutive bases of the polynucleotide of (a) or (b).

2. The polynucleotide of Claim 1 wherein the polynucleotide is DNA.

3. The polynucleotide of Claim 1 wherein the polynucleotide is RNA.

4. The polynucleotide of Claim 2 which encodes an enzyme comprising amino acids 1 to 414 of SEQ ID NO:33.

5. The polynucleotide of Claim 2 which encodes an enzyme comprising amino acids 1 to 373 of SEQ ID NO:34.

6. The polynucleotide of Claim 2 which encodes an enzyme comprising amino acids 1 to 453 of SEQ ID NO:35.

7. The polynucleotide of Claim 2 which encodes an enzyme comprising amino acids 1 to 343 of SEQ ID NO:36.

8. The polynucleotide of Claim 2 which encodes an enzyme comprising amino acids 1 to 398 of SEQ ID NO:37.

9. The polynucleotide of Claim 2 which encodes an enzyme comprising amino acids 1 to 592 of SEQ ID NO:38.

10. The polynucleotide of Claim 2 which encodes an enzyme comprising amino acids 1 to 354 of SEQ ID NO:39.

11. The polynucleotide of Claim 2 which encodes an enzyme comprising amino acids 1 to 303 of SEQ ID NO:40.

12. The polynucleotide of Claim 2 which encodes an enzyme comprising amino acids 1 to 311 of SEQ ID NO:41.

13. The polynucleotide of Claim 2 which encodes an enzyme comprising amino acids 1 to 305 of SEQ ID NO:42.

14. An isolated polynucleotide comprising a member selected from the group consisting of:

- (a) a polynucleotide having at least a 70% identity to a polynucleotide encoding an enzyme expressed by the DNA contained in ATCC Deposit No. \_\_\_\_\_;
- (b) a polynucleotide complementary to the polynucleotide of (a); and
- (c) a polynucleotide comprising at least 15 consecutive bases of the polynucleotide of (a) and (b).

15. A vector comprising the DNA of Claim 2.

16. A host cell comprising the vector of Claim 15.

17. A process for producing a polypeptide comprising: expressing from the host cell of Claim 16 a polypeptide encoded by said DNA.

18. A process for producing a cell comprising: transforming or transfecting the cell with the vector of Claim 15 such that the cell expresses the polypeptide encoded by the DNA contained in the vector.

19. An enzyme comprising a member selected from the group consisting of an enzyme comprising an amino acid sequence which is at least 70% identical to the amino acid sequence set forth in SEQ ID NOS:33-42.

20. A method for transferring an amino group from an amino acid to an  $\alpha$ -keto acid comprising:

contacting an amino acid in the presence of an  $\alpha$ -keto acid with an enzyme selected from the group consisting of an enzyme having the amino acid sequence set forth in SEQ ID NOS:33-42.

### **ABSTRACT**

Esterase enzymes derived from various *Staphylothermus*, *Pyrodictium*, *Archaeoglobus*, *Aquifex*, M11TL, *Thermococcus*, *Teredinibacter* and *Sulfolobus* organisms are disclosed. The enzymes are produced from native or recombinant host cells and can be utilized in the pharmaceutical, agricultural and other industries.



FIGURE 1

*Staphylothermus marinus* - F1-12LC

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ATG | TCT | TTA | AAC | AAG | CAC | TCT | TGG | ATG | GAT | ATG | ATA | ATA | TTT | ATT | CTC |
| Met | Ser | Leu | Asn | Lys | His | Ser | Trp | Met | Asp | Met | Ile | Ile | Phe | Ile | Leu |
| AGC | TTT | TCT | TTC | CCA | TTA | ACA | ATG | ATC | GCA | TTA | GCT | ATC | TCT | ATG | TCG |
| Ser | Phe | Ser | Phe | Pro | Leu | Thr | Met | Ile | Ala | Leu | Ala | Ile | Ser | Met | Ser |
| TCA | TGG | TTT | AAT | ATA | TGG | AAT | AAT | GCA | TTA | AGC | GAT | CTA | GGA | CAT | GCT |
| Ser | Trp | Phe | Asn | Ile | Trp | Asn | Asn | Ala | Leu | Ser | Asp | Leu | Gly | His | Ala |
| GTT | AAA | AGC | AGT | GTT | GCT | CCA | ATA | TTC | AAT | CTA | GGT | CTT | GCA | ATT | GGT |
| Val | Lys | Ser | Ser | Val | Ala | Pro | Ile | Phe | Asn | Leu | Gly | Leu | Ala | Ile | Gly |
| GGG | ATA | CTA | ATT | GTT | ATA | GTT | GGT | TTA | AGA | AAT | CTT | TAT | TCG | TGG | AGT |
| Gly | Ile | Leu | Ile | Val | Ile | Val | Gly | Leu | Arg | Asn | Leu | Tyr | Ser | Trp | Ser |
| AGA | GTT | AAA | GGA | TCT | TTA | ATC | ATA | TCC | ATG | GGT | GTA | TTT | CTT | AAC | TTA |
| Arg | Val | Lys | Gly | Ser | Leu | Ile | Ile | Ser | Met | Gly | Val | Phe | Leu | Asn | Leu |
| ATA | GGG | GTT | TTC | GAC | GAA | GTA | TAT | GGT | TGG | ATA | CAT | TTC | CTA | GTC | TCA |
| Ile | Gly | Val | Phe | Asp | Glu | Val | Tyr | Gly | Trp | Ile | His | Phe | Leu | Val | Ser |
| GTA | TTG | TTT | TTC | TTA | TCA | ATA | ATA | GCA | TAT | TTC | ATA | GCT | ATA | TCA | ATA |
| Val | Leu | Phe | Phe | Leu | Ser | Ile | Ile | Ala | Tyr | Phe | Ile | Ala | Ile | Ser | Ile |
| CTT | GAC | AAA | TCA | TGG | ATA | GCT | GTT | CTA | CTA | ATA | ATA | GGT | CAT | ATT | GCA |
| Leu | Asp | Lys | Ser | Trp | Ile | Ala | Val | Leu | Leu | Ile | Ile | Gly | His | Ile | Ala |
| ATG | TGG | TAT | CTA | CAC | TTT | GCT | TCA | GAG | ATT | CCG | AGA | GGT | GCG | GCT | ATT |
| Met | Trp | Tyr | Leu | His | Phe | Ala | Ser | Glu | Ile | Pro | Arg | Gly | Ala | Ala | Ile |
| CCC | GAG | TTA | TTA | GCG | GTA | TTC | TCG | TTT | TTA | CCA | TTC | TAT | ATA | AGA | GAC |
| Pro | Glu | Leu | Leu | Ala | Val | Phe | Ser | Phe | Leu | Pro | Phe | Tyr | Ile | Arg | Asp |
| TAT | TTT | AAA | TCA | TAC | ACT | AAA | CGA | TAG |     |     |     |     |     |     |     |
| Tyr | Phe | Lys | Ser | Tyr | Thr | Lys | Arg |     |     |     |     |     |     |     |     |

FIGURE 2

*Pyrodictium* - TAG11-17LC

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ATG | AAA | CTC | CTT | GAG | CCC | ACA | AAT | ACC | TCC | TAC | ACG | CTG | TTA | CAG | GAT |
| Met | Lys | Leu | Leu | Glu | Pro | Thr | Asn | Thr | Ser | Tyr | Thr | Leu | Leu | Gln | Asp |
| TTA | GCA | TTG | CAT | TTT | GCA | TTT | TAC | TGG | TTT | CTG | GCC | GTG | TAT | ACG | TGG |
| Leu | Ala | Leu | His | Phe | Ala | Phe | Tyr | Trp | Phe | Leu | Ala | Val | Tyr | Thr | Trp |
| TTA | CCC | GGT | GTC | CTA | GTC | CGG | GGC | GTA | GCT | GTG | GAC | ACA | GGG | GTG | GCT |
| Leu | Pro | Gly | Val | Leu | Val | Arg | Gly | Val | Ala | Val | Asp | Thr | Gly | Val | Ala |
| CGG | GTG | CCT | GGG | CTC | GGC | CGG | CGC | GGT | AAG | AGG | CTG | CTC | CTG | GCC | GCT |
| Arg | Val | Pro | Gly | Leu | Gly | Arg | Arg | Gly | Lys | Arg | Leu | Leu | Leu | Ala | Ala |
| GTG | GCT | GTC | TTG | GCG | CTT | GTT | GTG | TCC | GTT | GTT | GTC | CCG | GCT | TAT | GTG |
| Val | Ala | Val | Leu | Ala | Leu | Val | Val | Ser | Val | Val | Val | Pro | Ala | Tyr | Val |
| GCG | TAT | AGT | AGT | CTG | CAC | CCG | GAG | AGC | TGT | CGG | CCC | GTT | GCG | CCG | GAG |
| Ala | Tyr | Ser | Ser | Leu | His | Pro | Glu | Ser | Cys | Arg | Pro | Val | Ala | Pro | Glu |
| GGG | CTC | ACC | TAC | AAA | GAG | TTC | AGC | GTG | ACC | GCG | GAG | GAT | GGC | TTG | GTG |
| Gly | Leu | Thr | Tyr | Lys | Glu | Phe | Ser | Val | Thr | Ala | Glu | Asp | Gly | Leu | Val |
| GTT | CGG | GGC | TGG | GTG | CTG | GGC | CCC | GGC | GCT | GGG | GGC | AAC | CCG | GTG | TTC |
| Val | Arg | Gly | Trp | Val | Leu | Gly | Pro | Gly | Ala | Gly | Gly | Asn | Pro | Val | Phe |
| GTT | TTG | ATG | CAC | GGG | TAT | ACT | GGG | TGC | CGC | TCG | GCG | CCC | TAC | ATG | GCT |
| Val | Leu | Met | His | Gly | Tyr | Thr | Gly | Cys | Arg | Ser | Ala | Pro | Tyr | Met | Ala |
| GTG | CTG | GCC | CGG | GAG | CTC | GTG | GAG | TGG | GGG | TAC | CCG | GTG | GTT | GTG | TTC |
| Val | Leu | Ala | Arg | Glu | Leu | Val | Glu | Trp | Gly | Tyr | Pro | Val | Val | Val | Phe |
| GAC | TTC | CGG | GGC | CAC | GGG | GAG | AGC | GGG | GGC | TCG | ACG | ACG | ATT | GGG | CCC |
| Asp | Phe | Arg | Gly | His | Gly | Glu | Ser | Gly | Gly | Ser | Thr | Thr | Ile | Gly | Pro |
| CGG | GAG | GTG | CTG | GAT | GCC | CGG | GCT | GTG | GTG | GGC | TAT | GTC | TCG | GAG | CGG |
| Arg | Glu | Val | Leu | Asp | Ala | Arg | Ala | Val | Val | Gly | Tyr | Val | Ser | Glu | Arg |
| TTC | CCC | GGC | CGC | CGG | ATA | ATA | TTG | GTG | GGG | TTC | AGT | ATG | GGC | GGC | GCT |
| Phe | Pro | Gly | Arg | Arg | Ile | Ile | Leu | Val | Gly | Phe | Ser | Met | Gly | Gly | Ala |
| GTA | GCG | ATC | GTG | GAG | GGT | GCT | GGG | GAC | CCG | CGG | GTC | TAC | GCG | GTG | GCT |
| Val | Ala | Ile | Val | Glu | Gly | Ala | Gly | Asp | Pro | Arg | Val | Tyr | Ala | Val | Ala |
| GCT | GAT | AGC | CCG | TAC | TAT | AGG | CTC | CGG | GAC | GTC | ATA | CCC | CGG | TGG | CTG |
| Ala | Asp | Ser | Pro | Tyr | Tyr | Arg | Leu | Arg | Asp | Val | Ile | Pro | Arg | Trp | Leu |
| GAG | TAC | AAG | ACG | CCG | CTG | CCG | GGC | TGG | GTG | GGT | GTG | CTG | GCC | GGG | TTC |



FIGURE 3

*Archaeoglobus Venificus* SN P6-24LC

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ATG | CCA | TAT | GTT | AGG | AAT | GGT | GGT | GTA | AAT | ATC | TAT | TAT | GAA | CTG | GTG |
| Met | Pro | Tyr | Val | Arg | Asn | Gly | Gly | Val | Asn | Ile | Tyr | Tyr | Glu | Leu | Val |
| GAT | GGA | CCT | GAG | CCA | CCA | ATT | GTC | TTT | GTT | CAC | GGA | TGG | ACA | GCA | AAT |
| Asp | Gly | Pro | Glu | Pro | Pro | Ile | Val | Phe | Val | His | Gly | Trp | Thr | Ala | Asn |
| ATG | AAT | TTT | TGG | AAA | GAG | CAA | AGA | CGT | TAT | TTT | GCA | GGC | AGG | AAT | ATG |
| Met | Asn | Phe | Trp | Lys | Glu | Gln | Arg | Arg | Tyr | Phe | Ala | Gly | Arg | Asn | Met |
| ATG | TTG | TTT | GTC | GAT | AAC | AGA | GGT | CAT | GGC | AGG | TCC | GAT | AAG | CCA | CTT |
| Met | Leu | Phe | Val | Asp | Asn | Arg | Gly | His | Gly | Arg | Ser | Asp | Lys | Pro | Leu |
| GGA | TAC | GAT | TTC | TAC | AGA | TTT | GAG | AAC | TTC | ATT | TCA | GAT | TTA | GAT | GCG |
| Gly | Tyr | Asp | Phe | Tyr | Arg | Phe | Glu | Asn | Phe | Ile | Ser | Asp | Leu | Asp | Ala |
| GTT | GTT | AGG | GAG | ACT | GGA | GTG | GAG | AAA | TTT | GTT | CTC | GTC | GGA | CAT | TCA |
| Val | Val | Arg | Glu | Thr | Gly | Val | Glu | Lys | Phe | Val | Leu | Val | Gly | His | Ser |
| TTC | GGA | ACA | ATG | ATC | TCT | ATG | AAG | TAC | TGT | TCG | GAG | TAT | CGG | AAT | CGG |
| Phe | Gly | Thr | Met | Ile | Ser | Met | Lys | Tyr | Cys | Ser | Glu | Tyr | Arg | Asn | Arg |
| GTT | CTT | GCT | CTA | ATC | CTC | ATA | GGT | GGT | GGG | AGC | AGA | ATA | AAG | CTT | CTA |
| Val | Leu | Ala | Leu | Ile | Leu | Ile | Gly | Gly | Gly | Ser | Arg | Ile | Lys | Leu | Leu |
| CAC | AGA | ATT | GGA | TAT | CCT | TTA | GCA | AAG | ATT | CTT | GCA | TCC | ATT | GCA | TAC |
| His | Arg | Ile | Gly | Tyr | Pro | Leu | Ala | Lys | Ile | Leu | Ala | Ser | Ile | Ala | Tyr |
| AAG | AAG | TCT | TCA | AGA | TTG | GTC | GCA | GAT | CTT | TCC | TTT | GGC | AAA | AAT | GCT |
| Lys | Lys | Ser | Ser | Arg | Leu | Val | Ala | Asp | Leu | Ser | Phe | Gly | Lys | Asn | Ala |
| GGT | GAA | CTT | AAA | GAG | TGG | GGA | TGG | AAA | CAG | GCA | ATG | GAT | TAT | ACA | CCC |
| Gly | Glu | Leu | Lys | Glu | Trp | Gly | Trp | Lys | Gln | Ala | Met | Asp | Tyr | Thr | Pro |
| TCC | TAC | GTG | GCA | ATG | TAC | ACG | TAC | AGA | ACT | CTA | ACG | AAA | GTG | AAT | CTT |
| Ser | Tyr | Val | Ala | Met | Tyr | Thr | Tyr | Arg | Thr | Leu | Thr | Lys | Val | Asn | Leu |
| GAA | AAT | ATC | TTG | GAG | AAA | ATA | GAC | TGT | CCA | ACA | CTG | ATT | ATC | GTT | GGA |
| Glu | Asn | Ile | Leu | Glu | Lys | Ile | Asp | Cys | Pro | Thr | Leu | Ile | Ile | Val | Gly |
| GAA | GAG | GAT | GCA | CTA | TTG | CCC | GTT | AGC | AAA | TCA | GTT | GAG | CTG | AGC | AGG |
| Glu | Glu | Asp | Ala | Leu | Leu | Pro | Val | Ser | Lys | Ser | Val | Glu | Leu | Ser | Arg |
| AGG | ATA | GAA | AAC | TCA | AAG | CTT | GTG | ATC | ATC | CCA | AAC | TCG | GGG | CAT | TGC |
| Arg | Ile | Glu | Asn | Ser | Lys | Leu | Val | Ile | Ile | Pro | Asn | Ser | Gly | His | Cys |
| GTA | ATG | CTT | GAG | AGT | CCA | AGT | GAG | GTT | AAT | AGA | GCA | ATG | GAC | GAA | TTC |



FIGURE 4

*Aquifex pyrophilus* - 28LC

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| TTG | AGA | TTG | AGG | AAA | TTT | GAA | GAG | ATA | AAC | CTC | GTT | CTT | TCG | GGA | GGA |
| Leu | Arg | Leu | Arg | Lys | Phe | Glu | Glu | Ile | Asn | Leu | Val | Leu | Ser | Gly | Gly |
| GCT | GCA | AAG | GGC | ATA | GCC | CAC | ATA | GGT | GTT | TTG | AAA | GCT | ATA | AAC | GAG |
| Ala | Ala | Lys | Gly | Ile | Ala | His | Ile | Gly | Val | Leu | Lys | Ala | Ile | Asn | Glu |
| CTC | GGT | ATA | AGG | GTG | AGG | GCT | TTA | AGC | GGG | GTG | AGC | GCC | GGG | GCA | ATC |
| Leu | Gly | Ile | Arg | Val | Arg | Ala | Leu | Ser | Gly | Val | Ser | Ala | Gly | Ala | Ile |
| GTT | TCG | GTC | TTT | TAT | GCC | TCA | GGC | TAC | TCC | CCT | GAA | GGG | ATG | TTC | AGC |
| Val | Ser | Val | Phe | Tyr | Ala | Ser | Gly | Tyr | Ser | Pro | Glu | Gly | Met | Phe | Ser |
| CTT | CTG | AAG | AGG | GTA | AAC | TGG | CTG | AAG | CTG | TTT | AAG | TTC | AAG | CCA | CCT |
| Leu | Leu | Lys | Arg | Val | Asn | Trp | Leu | Lys | Leu | Phe | Lys | Phe | Lys | Pro | Pro |
| CTG | AAG | GGA | TTG | ATA | GGG | TGG | GAG | AAG | GCT | ATA | AGA | TTC | CTT | GAG | GAA |
| Leu | Lys | Gly | Leu | Ile | Gly | Trp | Glu | Lys | Ala | Ile | Arg | Phe | Leu | Glu | Glu |
| GTT | CTC | CCT | TAC | AGG | AGA | ATA | GAA | AAA | CTT | GAG | ATA | CCG | ACG | TAT | ATA |
| Val | Leu | Pro | Tyr | Arg | Arg | Ile | Glu | Lys | Leu | Glu | Ile | Pro | Thr | Tyr | Ile |
| TGC | GCG | ACG | GAT | TTA | TAC | TCG | GGA | AGG | GCT | CTA | TAC | CTC | TCG | GAA | GGG |
| Cys | Ala | Thr | Asp | Leu | Tyr | Ser | Gly | Arg | Ala | Leu | Tyr | Leu | Ser | Glu | Gly |
| AGT | TTA | ATC | CCC | GCA | CTT | CTC | GGC | AGC | TGT | GCA | ATT | CCC | GGC | ATA | TTT |
| Ser | Leu | Ile | Pro | Ala | Leu | Leu | Gly | Ser | Cys | Ala | Ile | Pro | Gly | Ile | Phe |
| GAA | CCC | GTT | GAG | TAT | AAG | AAT | TAC | TTG | CTC | GTT | GAC | GGA | GGT | ATA | GTT |
| Glu | Pro | Val | Glu | Tyr | Lys | Asn | Tyr | Leu | Leu | Val | Asp | Gly | Gly | Ile | Val |
| AAC | AAC | CTT | CCC | GTT | GAG | CCC | TTT | CAG | GAA | AGC | GGT | ATT | CCC | ACC | GTT |
| Asn | Asn | Leu | Pro | Val | Glu | Pro | Phe | Gln | Glu | Ser | Gly | Ile | Pro | Thr | Val |
| TGC | GTT | GAT | GTC | CTT | CCC | ATA | GAG | CCG | GAA | AAG | GAT | ATA | AAG | AAC | ATT |
| Cys | Val | Asp | Val | Leu | Pro | Ile | Glu | Pro | Glu | Lys | Asp | Ile | Lys | Asn | Ile |
| CTT | CAC | ATC | CTT | TTG | AGG | AGC | TTC | TTT | CTT | GCG | GTC | CGC | TCA | AAC | TCC |
| Leu | His | Ile | Leu | Leu | Arg | Ser | Phe | Phe | Leu | Ala | Val | Arg | Ser | Asn | Ser |
| GAA | AAG | AGA | AAG | GAG | TTT | TGT | GAC | CTC | GTT | ATA | GTT | CCT | GAG | CTT | GAG |
| Glu | Lys | Arg | Lys | Glu | Phe | Cys | Asp | Leu | Val | Ile | Val | Pro | Glu | Leu | Glu |
| GAG | TTC | ACA | CCC | CTT | GAT | GTT | AGA | AAA | GCG | GAC | CAA | ATA | ATG | GAG | AGG |
| Glu | Phe | Thr | Pro | Leu | Asp | Val | Arg | Lys | Ala | Asp | Gln | Ile | Met | Glu | Arg |



## FIGURE 5

## M11TL-29L.

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ATG | TTT | AAT | ATC | AAT | GTC | TTT | GTT | AAT | ATA | TCT | TGG | CTG | TAT | TTT | TCA |
| Met | Phe | Asn | Ile | Asn | Val | Phe | Val | Asn | Ile | Ser | Trp | Leu | Tyr | Phe | Ser |
| GGG | ATA | GTT | ATG | AAG | ACT | GTG | GAA | GAG | TAT | GCG | CTA | CTT | GAA | ACA | GGC |
| Gly | Ile | Val | Met | Lys | Thr | Val | Glu | Glu | Tyr | Ala | Leu | Leu | Glu | Thr | Gly |
| GTA | AGA | GTG | TTT | TAT | CGG | TGT | GTA | ATC | CCG | GAG | AAA | GCT | TTT | AAC | ACT |
| Val | Arg | Val | Phe | Tyr | Arg | Cys | Val | Ile | Pro | Glu | Lys | Ala | Phe | Asn | Thr |
| TTG | ATA | ATA | GGT | TCA | CAC | GGA | TTG | GGG | GCG | CAC | AGT | GGA | ATC | TAC | ATT |
| Leu | Ile | Ile | Gly | Ser | His | Gly | Leu | Gly | Ala | His | Ser | Gly | Ile | Tyr | Ile |
| AGT | GTT | GCT | GAA | GAA | TTT | GCT | AGG | CAC | GGA | TTT | GGA | TTC | TGC | ATG | CAC |
| Ser | Val | Ala | Glu | Glu | Phe | Ala | Arg | His | Gly | Phe | Gly | Phe | Cys | Met | His |
| GAT | CAA | AGG | GGA | CAT | GGG | AGA | ACG | GCA | AGC | GAT | AGA | GAA | AGA | GGG | TAT |
| Asp | Gln | Arg | Gly | His | Gly | Arg | Thr | Ala | Ser | Asp | Arg | Glu | Arg | Gly | Tyr |
| GTG | GAG | GGC | TTT | CAC | AAC | TTC | ATA | GAG | GAT | ATG | AAG | GCC | TTC | TCC | GAT |
| Val | Glu | Gly | Phe | His | Asn | Phe | Ile | Glu | Asp | Met | Lys | Ala | Phe | Ser | Asp |
| TAT | GCC | AAG | TGG | CGC | GTG | GGA | GGT | GAC | GAA | ATA | ATA | TTG | CTA | GGA | CAC |
| Tyr | Ala | Lys | Trp | Arg | Val | Gly | Gly | Asp | Glu | Ile | Ile | Leu | Leu | Gly | His |
| AGT | ATG | GGC | GGG | CTG | ATA | GCG | CTC | TTA | ACA | GTT | GCA | ACT | TAT | AAA | GAA |
| Ser | Met | Gly | Gly | Leu | Ile | Ala | Leu | Leu | Thr | Val | Ala | Thr | Tyr | Lys | Glu |
| ATC | GCC | AAG | GGA | GTT | ATC | GCG | CTA | GCC | CCG | GCC | CTC | CAA | ATC | CCC | TTA |
| Ile | Ala | Lys | Gly | Val | Ile | Ala | Leu | Ala | Pro | Ala | Leu | Gln | Ile | Pro | Leu |
| ACC | CCG | GCT | AGA | AGA | CTT | GTT | CTA | AGC | CTC | GCG | TCA | AGG | CTT | GCC | CCG |
| Thr | Pro | Ala | Arg | Arg | Leu | Val | Leu | Ser | Leu | Ala | Ser | Arg | Leu | Ala | Pro |
| CAT | TCT | AAG | ATC | ACC | TTA | CAA | AGG | AGA | TTG | CCG | CAG | AAA | CCA | GAG | GGT |
| His | Ser | Lys | Ile | Thr | Leu | Gln | Arg | Arg | Leu | Pro | Gln | Lys | Pro | Glu | Gly |
| TTT | CAA | AGA | GCA | AAA | GAT | ATA | GAA | TAC | AGT | CTG | AGT | GAA | ATA | TCA | GTC |
| Phe | Gln | Arg | Ala | Lys | Asp | Ile | Glu | Tyr | Ser | Leu | Ser | Glu | Ile | Ser | Val |
| AAG | CTC | GTG | GAC | GAA | ATG | ATT | AAA | GCA | TCA | TCT | ATG | TTC | TGG | ACC | ATA |
| Lys | Leu | Val | Asp | Glu | Met | Ile | Lys | Ala | Ser | Ser | Met | Phe | Trp | Thr | Ile |
| GCA | GGG | GAA | ATT | AAT | ACT | CCC | GTC | CTG | CTT | ATT | CAT | GGG | GAA | AAA | GAC |
| Ala | Gly | Glu | Ile | Asn | Thr | Pro | Val | Leu | Leu | Ile | His | Gly | Glu | Lys | Asp |



AAT GTC ATA CCT CCG GAG GCG AGC AAA AAA GCC TAC CAA TTA ATA CCT  
Asn Val Ile Pro Pro Glu Ala Ser Lys Lys Ala Tyr Gln Leu Ile Pro

8/33

TCA TTC CCT AAA GAG TTG AAA ATA TAC CCC GAT CTT GGA CAC AAC TTG  
Ser Phe Pro Lys Glu Leu Lys Ile Tyr Pro Asp Leu Gly His Asn Leu

TTT TTT GAA CCA GGC GCG GTG AAA ATC GTC ACA GAC ATT GTA GAG TGG  
Phe Phe Glu Pro Gly Ala Val Lys Ile Val Thr Asp Ile Val Glu Trp

GTT AAG AAT CTA CCC AGG GAA AAT CCT TAA  
Val Lys Asn Leu Pro Arg Glu Asn Pro

6644330" 24233333

FIGURE 6

*Thermococcus* CL-2-30LC

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ATG | GAG | GTT | TAC | AAG | GCC | AAA | TTC | GGC | GAA | GCA | AAG | CTC | GGC | TGG | GTC |
| Met | Glu | Val | Tyr | Lys | Ala | Lys | Phe | Gly | Glu | Ala | Lys | Leu | Gly | Trp | Val |
| GTT | CTG | GTT | CAT | GGC | CTC | GGC | GAG | CAC | AGC | GGA | AGG | TAT | GGA | AGA | CTG |
| Val | Leu | Val | His | Gly | Leu | Gly | Glu | His | Ser | Gly | Arg | Tyr | Gly | Arg | Leu |
| ATT | AAG | GAA | CTC | AAC | TAT | GCC | GGC | TTT | GGA | GTT | TAC | ACC | TTC | GAC | TGG |
| Ile | Lys | Glu | Leu | Asn | Tyr | Ala | Gly | Phe | Gly | Val | Tyr | Thr | Phe | Asp | Trp |
| CCC | GGC | CAC | GGG | AAG | AGC | CCG | GGC | AAG | AGA | GGG | CAC | ACG | AGC | GTC | GAG |
| Pro | Gly | His | Gly | Lys | Ser | Pro | Gly | Lys | Arg | Gly | His | Thr | Ser | Val | Glu |
| GAG | GCG | ATG | GAA | ATC | ATC | GAC | TCG | ATA | ATC | GAG | GAG | ATC | AGG | GAG | AAG |
| Glu | Ala | Met | Glu | Ile | Ile | Asp | Ser | Ile | Ile | Glu | Glu | Ile | Arg | Glu | Lys |
| CCC | TTC | CTC | TTC | GGC | CAC | AGC | CTC | GGT | GGT | CTA | ACT | GTC | ATC | AGG | TAC |
| Pro | Phe | Leu | Phe | Gly | His | Ser | Leu | Gly | Gly | Leu | Thr | Val | Ile | Arg | Tyr |
| GCT | GAG | ACG | CGG | CCC | GAT | AAA | ATA | CGG | GGA | TTA | ATA | GCT | TCC | TCG | CCT |
| Ala | Glu | Thr | Arg | Pro | Asp | Lys | Ile | Arg | Gly | Leu | Ile | Ala | Ser | Ser | Pro |
| GCC | CTC | GCC | AAG | AGC | CCG | GAA | ACG | CCG | GGC | TTC | ATG | GTG | GCC | CTC | GCG |
| Ala | Leu | Ala | Lys | Ser | Pro | Glu | Thr | Pro | Gly | Phe | Met | Val | Ala | Leu | Ala |
| AAG | TTC | CTT | GGA | AAG | ATC | GCC | CCG | GGA | GTT | GTT | CTC | TCC | AAC | GGC | ATA |
| Lys | Phe | Leu | Gly | Lys | Ile | Ala | Pro | Gly | Val | Val | Leu | Ser | Asn | Gly | Ile |
| AAG | CCG | GAA | CTC | CTC | TCG | AGG | AAC | AGG | GAC | GCC | GTG | AGG | AGG | TAC | GTT |
| Lys | Pro | Glu | Leu | Leu | Ser | Arg | Asn | Arg | Asp | Ala | Val | Arg | Arg | Tyr | Val |
| GAA | GAC | CCA | CTC | GTC | CAC | GAC | AGG | ATT | TCG | GCC | AAG | CTG | GGA | AGG | AGC |
| Glu | Asp | Pro | Leu | Val | His | Asp | Arg | Ile | Ser | Ala | Lys | Leu | Gly | Arg | Ser |
| ATC | TTC | GTG | AAC | ATG | GAG | CTG | GCC | CAC | AGG | GAG | GCG | GAC | AAG | ATA | AAA |
| Ile | Phe | Val | Asn | Met | Glu | Leu | Ala | His | Arg | Glu | Ala | Asp | Lys | Ile | Lys |
| GTC | CCG | ATC | CTC | CTT | CTG | ATC | GGC | ACT | GGC | GAT | GTA | ATA | ACC | CCG | CCT |
| Val | Pro | Ile | Leu | Leu | Leu | Ile | Gly | Thr | Gly | Asp | Val | Ile | Thr | Pro | Pro |
| GAA | GGC | TCA | CGC | AGA | CTC | TTC | GAG | GAG | CTG | GCC | GTC | GAG | AAC | AAA | ACC |
| Glu | Gly | Ser | Arg | Arg | Leu | Phe | Glu | Glu | Leu | Ala | Val | Glu | Asn | Lys | Thr |
| CTG | AGG | GAG | TTC | GAG | GGG | GCG | TAC | CAC | GAG | ATA | TTT | GAA | GAC | CCC | GAG |
| Leu | Arg | Glu | Phe | Glu | Gly | Ala | Tyr | His | Glu | Ile | Phe | Glu | Asp | Pro | Glu |



FIGURE 7

*Aquifex* VF5-34LC

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| TTG | ATT | GGC | AAT | TTG | AAA | TTG | AAG | AGG | TTT | GAA | GAG | GTT | AAC | TTA | GTT |  |
| Leu | Ile | Gly | Asn | Leu | Lys | Leu | Lys | Arg | Phe | Glu | Glu | Val | Asn | Leu | Val |  |
| CTT | TCG | GGA | GGG | GCT | GCC | AAG | GGT | ATC | GCC | CAT | ATA | GGT | GTT | TTA | AAA |  |
| Leu | Ser | Gly | Gly | Ala | Ala | Lys | Gly | Ile | Ala | His | Ile | Gly | Val | Leu | Lys |  |
| GCT | CTG | GAA | GAG | CTC | GGT | ATA | AAG | GTA | AAG | AGG | CTC | AGC | GGG | GTA | AGT |  |
| Ala | Leu | Glu | Glu | Leu | Gly | Ile | Lys | Val | Lys | Arg | Leu | Ser | Gly | Val | Ser |  |
| GCT | GGA | GCT | ATC | GTT | TCC | GTC | TTT | TAC | GCT | TCG | GGC | TAC | ACT | CCC | GAC |  |
| Ala | Gly | Ala | Ile | Val | Ser | Val | Phe | Tyr | Ala | Ser | Gly | Tyr | Thr | Pro | Asp |  |
| GAG | ATG | TTA | AAA | CTC | CTG | AAA | GAG | GTA | AAC | TGG | CTC | AAA | CTT | TTT | AAG |  |
| Glu | Met | Leu | Lys | Leu | Leu | Lys | Glu | Val | Asn | Trp | Leu | Lys | Leu | Phe | Lys |  |
| TTC | AAA | ACA | CCG | AAA | ATG | GGC | TTA | ATG | GGG | TGG | GAG | AAG | GCT | GCA | GAG |  |
| Phe | Lys | Thr | Pro | Lys | Met | Gly | Leu | Met | Gly | Trp | Glu | Lys | Ala | Ala | Glu |  |
| TTT | TTG | GAA | AAA | GAG | CTC | GGA | GTT | AAG | AGG | CTG | GAA | GAC | CTG | AAC | ATA |  |
| Phe | Leu | Glu | Lys | Glu | Leu | Gly | Val | Lys | Arg | Leu | Glu | Asp | Leu | Asn | Ile |  |
| CCA | ACC | TAT | CTT | TGC | TCG | GCG | GAT | CTG | TAC | ACG | GGA | AAG | GCT | CTT | TAC |  |
| Pro | Thr | Tyr | Leu | Cys | Ser | Ala | Asp | Leu | Tyr | Thr | Gly | Lys | Ala | Leu | Tyr |  |
| TTC | GGC | AGA | GGT | GAC | TTA | ATT | CCC | GTG | CTT | CTC | GGA | AGT | TGT | TCC | ATA |  |
| Phe | Gly | Arg | Gly | Asp | Leu | Ile | Pro | Val | Leu | Leu | Gly | Ser | Cys | Ser | Ile |  |
| CCC | GGG | ATT | TTT | GAA | CCA | GTT | GAG | TAC | GAG | AAT | TTT | CTA | CTT | GTT | GAC |  |
| Pro | Gly | Ile | Phe | Glu | Pro | Val | Glu | Tyr | Glu | Asn | Phe | Leu | Leu | Val | Asp |  |
| GGA | GGT | ATA | GTG | AAC | AAC | CTG | CCC | GTA | GAA | CCT | TTG | GAA | AAG | TTC | AAA |  |
| Gly | Gly | Ile | Val | Asn | Asn | Leu | Pro | Val | Glu | Pro | Leu | Glu | Lys | Phe | Lys |  |
| GAA | CCC | ATA | ATC | GGG | GTA | GAT | GTG | CTT | CCC | ATA | ACT | CAA | GAA | AGA | AAG |  |
| Glu | Pro | Ile | Ile | Gly | Val | Asp | Val | Leu | Pro | Ile | Thr | Gln | Glu | Arg | Lys |  |
| ATT | AAA | AAT | ATA | CTC | CAC | ATC | CTT | ATA | AGG | AGC | TTC | TTT | CTG | GCG | GTT |  |
| Ile | Lys | Asn | Ile | Leu | His | Ile | Leu | Ile | Arg | Ser | Phe | Phe | Leu | Ala | Val |  |
| CGT | TCC | AAT | TCG | GAA | AAG | AGA | AAG | GAG | TTC | TGC | AAC | GTA | GTT | ATA | GAA |  |
| Arg | Ser | Asn | Ser | Glu | Lys | Arg | Lys | Glu | Phe | Cys | Asn | Val | Val | Ile | Glu |  |
| CCT | CCC | CTT | GAA | GAG | TTC | TCT | CCT | CTG | GAC | GTA | AAT | AAG | GCG | GAC | GAG |  |
| Pro | Pro | Leu | Glu | Glu | Phe | Ser | Pro | Leu | Asp | Val | Asn | Lys | Ala | Asp | Glu |  |

ATA TTC TGC GGG GAT ATG AGA GCA CTT TAA  
Ile Phe Cys Gly Asp Met Arg Ala Leu

[illegible]

## FIGURE 8

*Teredinibacter* - 42L

ATG CCA GCT AAT GAC TCA CCC ACG ATC GAC TTT AAT CCT CGC GGC ATT  
 Met Pro Ala Asn Asp Ser Pro Thr Ile Asp Phe Asn Pro Arg Gly Ile  
  
 CTT CGC AAC GCT CAC GCA CAG GTT ATT TTA GCG ACT TCC GGC TTG CGC  
 Leu Arg Asn Ala His Ala Gln Val Ile Leu Ala Thr Ser Gly Leu Arg  
  
 AAA GCG TTT TTG AAA CGC ACG CAC AAG AGC TAC CTC AGC ACT GCC CAA  
 Lys Ala Phe Leu Lys Arg Thr His Lys Ser Tyr Leu Ser Thr Ala Gln  
  
 TGG CTG GAG CTC GAT GCC GGC AAC GGA GTT ACC TTG GCC GGA GAG CTT  
 Trp Leu Glu Leu Asp Ala Gly Asn Gly Val Thr Leu Ala Gly Glu Leu  
  
 AAC ACA GCG CCT GCA ACT GCA TCC TCC TCC CAC CCG GCG CAC AAG AAC  
 Asn Thr Ala Pro Ala Thr Ala Ser Ser Ser His Pro Ala His Lys Asn  
  
 ACT CTG GTT ATT GTG CTG CAC GGC TGG GAA GGC TCC AGC CAG TCG GCC  
 Thr Leu Val Ile Val Leu His Gly Trp Glu Gly Ser Ser Gln Ser Ala  
  
 TAT GCG ACC TCC GCT GGC AGC ACG CTT TTC GAC AAT GGG TTC GAC ACT  
 Tyr Ala Thr Ser Ala Gly Ser Thr Leu Phe Asp Asn Gly Phe Asp Thr  
  
 TTT CGC CTT AAT TTT CGC GAT CAC GGC GAC ACC TAC CAC TTA AAC CGC  
 Phe Arg Leu Asn Phe Arg Asp His Gly Asp Thr Tyr His Leu Asn Arg  
  
 GGC ATA TTT AAC TCA TCG CTG ATT GAC GAA GTA GTG GGC GCA GTC AAA  
 Gly Ile Phe Asn Ser Ser Leu Ile Asp Glu Val Val Gly Ala Val Lys  
  
 GCC ATC CAG CAG CAA ACC GAC TAC GAC AAG TAT TGC CTG ATG GGG TTC  
 Ala Ile Gln Gln Gln Thr Asp Tyr Asp Lys Tyr Cys Leu Met Gly Phe  
  
 TCA CTG GGT GGG AAC TTT GCC TTG CGC GTC GCG GTG CGG GAA CAG CAT  
 Ser Leu Gly Gly Asn Phe Ala Leu Arg Val Ala Val Arg Glu Gln His  
  
 CTC GCT AAA CCG CTA GCG GGC GTG CTC GCC GTA TGC CCG GTA CTC GAC  
 Leu Ala Lys Pro Leu Ala Gly Val Leu Ala Val Cys Pro Val Leu Asp  
  
 CCC GCA CAC ACC ATG ATG GCC CTA AAC CGA GGT GCG TTT TTC TAC GGC  
 Pro Ala His Thr Met Met Ala Leu Asn Arg Gly Ala Phe Phe Tyr Gly  
  
 CGC TAT TTT GCG CAT AAA TGG AAG CGC TCG TTA ACC GCA AAA CTT GCA  
 Arg Tyr Phe Ala His Lys Trp Lys Arg Ser Leu Thr Ala Lys Leu Ala  
  
 GCT TTC CCA GAC TAC AAA TAC GGC AAA GAT TTA AAA TCG ATA CAC ACG  
 Ala Phe Pro Asp Tyr Lys Tyr Gly Lys Asp Leu Lys Ser Ile His Thr



## FIGURE 9

*Archeoglobus fulgidus* VC16 - 16MC1

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ATG | CTT | GAT | ATG | CCA | ATC | GAC | CCT | GTT | TAC | TAC | CAG | CTT | GCT | GAG | TAT |
| Met | Leu | Asp | Met | Pro | Ile | Asp | Pro | Val | Tyr | Tyr | Gln | Leu | Ala | Glu | Tyr |
| TTC | GAC | AGT | CTG | CCG | AAG | TTC | GAC | CAG | TTT | TCC | TCG | GCC | AGA | GAG | TAC |
| Phe | Asp | Ser | Leu | Pro | Lys | Phe | Asp | Gln | Phe | Ser | Ser | Ala | Arg | Glu | Tyr |
| AGG | GAG | GCG | ATA | AAT | CGA | ATA | TAC | GAG | GAG | AGA | AAC | CGG | CAG | CTG | AGC |
| Arg | Glu | Ala | Ile | Asn | Arg | Ile | Tyr | Glu | Glu | Arg | Asn | Arg | Gln | Leu | Ser |
| CAG | CAT | GAG | AGG | GTT | GAA | AGA | GTT | GAG | GAC | AGG | ACG | ATT | AAG | GGG | AGG |
| Gln | His | Glu | Arg | Val | Glu | Arg | Val | Glu | Asp | Arg | Thr | Ile | Lys | Gly | Arg |
| AAC | GGA | GAC | ATC | AGA | GTC | AGA | GTT | TAC | CAG | CAG | AAG | CCC | GAT | TCC | CCG |
| Asn | Gly | Asp | Ile | Arg | Val | Arg | Val | Tyr | Gln | Gln | Lys | Pro | Asp | Ser | Pro |
| GGT | CTG | GTT | TAC | TAT | CAC | GGT | GGT | GGA | TTT | GTG | ATT | TGC | AGC | ATC | GAG |
| Val | Leu | Val | Tyr | Tyr | His | Gly | Gly | Gly | Phe | Val | Ile | Cys | Ser | Ile | Glu |
| TCG | CAC | GAC | GCC | TTA | TGC | AGG | AGA | AYY | GCG | AGA | CTT | TCA | AAC | TCT | ACC |
| Ser | His | Asp | Ala | Leu | Cys | Arg | Arg | Ile | Ala | Arg | Leu | Ser | Asn | Ser | Thr |
| GTA | GTC | TCC | GTG | GAT | TAC | AGG | CTC | GCT | CCT | GAG | CAC | AAG | TTT | CCC | CCC |
| Val | Val | Ser | Val | Asp | Tyr | Arg | Leu | Ala | Pro | Glu | His | Lys | Phe | Pro | Ala |
| CCA | GTT | TAT | CAT | TGC | TAC | GAT | GCG | ACC | AAG | TGG | GTT | GCT | GAG | AAC | CGG |
| Ala | Val | Tyr | Asp | Cys | Tyr | Asp | Ala | Thr | Lys | Trp | Val | Ala | Glu | Asn | Ala |
| GAG | GAG | CTG | AGG | ATT | GAC | CCG | TCA | AAA | ATC | TTC | GTT | GGG | GGG | GAC | AGT |
| Glu | Glu | Leu | Arg | Ile | Asp | Pro | Ser | Lys | Ile | Phe | Val | Gly | Gly | Asp | Ser |
| GCG | GGA | CGG | AAT | CTT | GCC | CCG | GCG | CTT | TCA | ATA | ATG | GCG | AGA | GAC | AGC |
| Ala | Gly | Gly | Asn | Leu | Ala | Ala | Ala | Val | Ser | Ile | Met | Ala | Arg | Asp | Ser |
| GGA | GAA | GAT | TTC | ATA | AAG | CAT | CAA | ATT | CTA | ACT | TAC | CCC | GTT | GTG | AAC |
| Gly | Glu | Asp | Phe | Ile | Lys | His | Gln | Ile | Leu | Ile | Tyr | Pro | Val | Val | Asn |
| TTT | GTA | GCC | CCC | ACA | CCA | TCG | CTT | CTG | GAG | TTT | GGA | GAG | GGG | CTG | TGG |
| Phe | Val | Ala | Pro | Thr | Pro | Ser | Leu | Leu | Glu | Phe | Gly | Glu | Gly | Leu | Trp |
| ATT | CTC | GAC | CAG | AAG | ATA | ATG | AGT | TGG | TTC | TCG | GAG | CAG | TAC | TTC | TCC |
| Ile | Leu | Asp | Gln | Lys | Ile | Met | Ser | Trp | Phe | Ser | Glu | Gln | Tyr | Phe | Ser |
| AGA | GAG | GAA | GAT | AAG | TTC | AAG | CCC | CTC | GCC | TCC | GTA | ATC | TTT | GCG | GAC |
| Arg | Glu | Glu | Asp | Lys | Phe | Asn | Pro | Leu | Ala | Ser | Val | Ile | Phe | Ala | Asp |





FIGURE 10

*Sulfolobus Solfataricus* P1 - 8LC1

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ATG | CCC | CTA | GAT | CCT | AGA | ATT | AAA | AAG | TTA | CTA | GAA | TCA | GCT | CTT | ACT |
| Met | Pro | Leu | Asp | Pro | Arg | Ile | Lys | Lys | Leu | Leu | Glu | Ser | Ala | Leu | Thr |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ATA | CCA | ATT | GGT | AAA | GCC | CCA | GTA | GAA | GAG | GTA | AGA | AAG | ATA | TTT | AGG |
| Ile | Pro | Ile | Gly | Lys | Ala | Pro | Val | Glu | Glu | Val | Arg | Lys | Ile | Phe | Arg |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| CAA | TTA | GCG | TCG | GCA | GCT | CCC | AAA | GTC | GAA | GTT | GGA | AAA | GTA | GAA | GAT |
| Gln | Leu | Ala | Ser | Ala | Ala | Pro | Lys | Val | Glu | Val | Gly | Lys | Val | Glu | Asp |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ATA | AAA | ATA | CCA | GGC | AGT | GAA | ACC | GTT | ATA | AAC | GCT | AGA | GTG | TAT | TTT |
| Ile | Lys | Ile | Pro | Gly | Ser | Glu | Thr | Val | Ile | Asn | Ala | Arg | Val | Tyr | Phe |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| CCG | AAG | AGT | AGC | GGT | CCT | TAT | GGT | GTT | CTA | GTG | TAT | CTT | CAT | GGA | GGC |
| Pro | Lys | Ser | Ser | Gly | Pro | Tyr | Gly | Val | Leu | Val | Tyr | Leu | His | Gly | Gly |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| GGT | TTT | GTA | ATA | GGC | GAT | GTG | GAA | TCT | TAT | GAC | CCA | TTA | TGT | AGA | GCA |
| Gly | Phe | Val | Ile | Gly | Asp | Val | Glu | Ser | Tyr | Asp | Pro | Leu | Cys | Arg | Ala |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ATT | ACA | AAT | GCG | TGC | AAT | TGC | GTT | GTA | GTA | TCA | GTG | GAC | TAT | AGG | TTA |
| Ile | Thr | Asn | Ala | Cys | Asn | Cys | Val | Val | Val | Ser | Val | Asp | Tyr | Arg | Leu |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| GCT | CCA | GAA | TAC | AAG | TTT | CCT | TCT | GCA | GTT | ATC | GAT | TCA | TTT | GAC | GCT |
| Ala | Pro | Glu | Tyr | Lys | Phe | Pro | Ser | Ala | Val | Ile | Asp | Ser | Phe | Asp | Ala |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ACT | AAT | TGG | GTT | TAT | AAC | AAT | TTA | GAT | AAA | TTT | GAT | GGA | AAG | ATG | GGA |
| Thr | Asn | Trp | Val | Tyr | Asn | Asn | Leu | Asp | Lys | Phe | Asp | Gly | Lys | Met | Gly |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| GTT | GCG | ATT | GCG | GGA | GAT | AGT | GCT | GGA | GGA | AAT | TTG | GCA | GCG | GTT | GTA |
| Val | Ala | Ile | Ala | Gly | Asp | Ser | Ala | Gly | Gly | Asn | Leu | Ala | Ala | Val | Val |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| GCT | CTT | CTT | TCA | AAG | GGT | AAA | ATT | AAT | TTG | AAG | TAT | CAA | ATA | CTG | GTT |
| Ala | Leu | Leu | Ser | Lys | Gly | Lys | Ile | Asn | Leu | Lys | Tyr | Gln | Ile | Leu | Val |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| TAC | CCA | GCG | GTA | AGT | TTA | GAT | AAC | GTT | TCA | AGA | TCC | ATG | ATA | GAG | TAC |
| Tyr | Pro | Ala | Val | Ser | Leu | Asp | Asn | Val | Ser | Arg | Ser | Met | Ile | Glu | Tyr |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| TCT | GAT | GGG | TTC | TTC | CTT | ACC | AGA | GAG | CAT | ATA | GAG | TGG | TTC | GGT | TCT |
| Ser | Asp | Gly | Phe | Phe | Leu | Thr | Arg | Glu | His | Ile | Glu | Trp | Phe | Gly | Ser |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| CAA | TAC | TTA | CGA | AGC | CCT | GCA | GAT | TTG | CTA | GAC | TTT | AGG | TTC | TCT | CCA |
| Gln | Tyr | Leu | Arg | Ser | Pro | Ala | Asp | Leu | Leu | Asp | Phe | Arg | Phe | Ser | Pro |



**Figure 11**  
**LA11.1 Esterase es2**

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ATG | AAG | GTT | AAA | CAC | GTT | ATT | GTT | TTA | CAT | GGC | TTA | TAT | ATG | TCT | GGC |
| Met | Lys | Val | Lys | His | Val | Ile | Val | Leu | His | Gly | Leu | Tyr | Met | Ser | Gly |
| TTG | GTG | ATG | CGC | CCG | TTA | TGT | TCG | CGT | CTA | GAA | GAG | TCG | GGG | GTT | AAA |
| Leu | Val | Met | Arg | Pro | Leu | Cys | Ser | Arg | Leu | Glu | Glu | Ser | Gly | Val | Lys |
| GTT | TTA | AAC | TTA | ACC | TAC | AAT | ACT | CGA | GAC | CCT | AAT | CGA | GAT | GCT | ATT |
| Val | Leu | Asn | Leu | Thr | Tyr | Asn | Thr | Arg | Asp | Pro | Asn | Arg | Asp | Ala | Ile |
| TTT | ACG | CAA | ATA | GAT | GAG | TTT | ATT | AGC | AAT | GAG | CCT | TCT | GCT | TTA | GTG |
| Phe | Thr | Gln | Ile | Asp | Glu | Phe | Ile | Ser | Asn | Glu | Pro | Ser | Ala | Leu | Val |
| TGT | CAC | TCT | ATG | GGG | GGC | TTA | GTT | GCT | CGC | GCC | TAT | TTA | GAG | GCA | AAC |
| Cys | His | Ser | Met | Gly | Gly | Leu | Val | Ala | Arg | Ala | Tyr | Leu | Glu | Ala | Asn |
| TCA | GCG | CCA | AGT | CAT | CAT | GTT | GAA | AAG | GTA | ATC | ACC | TTA | GGA | ACG | CCA |
| Ser | Ala | Pro | Ser | His | His | Val | Glu | Lys | Val | Ile | Thr | Leu | Gly | Thr | Pro |
| CAT | ACT | GGC | AGC | CAT | ATT | GCT | GAA | AAA | ATG | CAG | CAA | AAA | GGG | TTC | GAG |
| His | Thr | Gly | Ser | His | Ile | Ala | Glu | Lys | Met | Gln | Gln | Lys | Gly | Phe | Glu |
| CTA | TTA | TTA | AAA | AAT | AGC | GTT | GAG | TTT | TTA | CTC | TCT | AAG | AAT | GGT | GAT |
| Leu | Leu | Leu | Lys | Asn | Ser | Val | Glu | Phe | Leu | Leu | Ser | Lys | Asn | Gly | Asp |
| TGG | CCT | TTT | AAA | GCC | AAG | CTA | TAT | AGC | ATT | GCC | GGC | GAC | TTA | CCG | ATT |
| Trp | Pro | Phe | Lys | Ala | Lys | Leu | Tyr | Ser | Ile | Ala | Gly | Asp | Leu | Pro | Ile |
| GGC | TTA | ATG | CCA | CTC | ATT | GTA | AAA | GGC | AGC | CGC | TCT | GAT | GGC | ACT | GTA |
| Gly | Leu | Met | Pro | Leu | Ile | Val | Lys | Gly | Ser | Arg | Ser | Asp | Gly | Thr | Val |
| TTG | CTA | GAT | GAA | ACC | AAG | CTA | AAG | GGT | ATG | GCT | GAA | CAC | AAG | GTG | TTT |
| Leu | Leu | Asp | Glu | Thr | Lys | Leu | Lys | Gly | Met | Ala | Glu | His | Lys | Val | Phe |
| CAT | TTA | AGC | CAT | ACA | AGT | ATG | ATT | TAC | TCT | CGC | CAA | GTC | GTT | AAT | TAT |
| His | Leu | Ser | His | Thr | Ser | Met | Ile | Tyr | Ser | Arg | Gln | Val | Val | Asn | Tyr |
| ATT | CTT | GAG | CGC | TTG | AAC | GAG | GAC | ATT | TA  |     |     |     |     |     |     |
| Ile | Leu | Glu | Arg | Leu | Asn | Glu | Asp | Ile |     |     |     |     |     |     |     |



**Figure 13**  
**Metallosphaera Prunae Ron 12/2 Esterase 23mc1**

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ATG | CCC | CTA | CAT | CCA | AAG | GTA | AAG | AAA | TTA | CTT | TCC | CAG | CTA | CCT | CCC |
| Met | Pro | Leu | His | Pro | Lys | Val | Lys | Lys | Leu | Leu | Ser | Gln | Leu | Pro | Pro |
| CAG | GAC | TTC | TCC | AGA | AAC | GTG | CAG | GAC | CTG | AGG | AAG | GCC | TGG | GAT | TTA |
| Gln | Asp | Phe | Ser | Arg | Asn | Val | Gln | Asp | Leu | Arg | Lys | Ala | Trp | Asp | Leu |
| CCC | TTC | TCA | GGG | AGG | AGG | GAG | ACC | CTG | AAG | AGG | GTT | GAG | GAC | CTT | GAG |
| Pro | Phe | Ser | Gly | Arg | Arg | Glu | Thr | Leu | Lys | Arg | Val | Glu | Asp | Leu | Glu |
| ATA | CCC | ACT | AGG | GAC | GCA | CGA | ATC | AGG | GCC | AGG | GTC | TAC | ACC | CCC | TCA |
| Ile | Pro | Thr | Arg | Asp | Ala | Arg | Ile | Arg | Ala | Arg | Val | Tyr | Thr | Pro | Ser |
| AGT | AAG | GAA | AAC | TTA | CCC | GTC | CTT | GTT | TAC | TAT | CAC | GGC | GGT | GGC | TTC |
| Ser | Lys | Glu | Asn | Leu | Pro | Val | Leu | Val | Tyr | Tyr | His | Gly | Gly | Gly | Phe |
| GTG | TTC | GGT | AGC | GTT | GAC | AGC | TAC | GAC | GGC | CTC | GCA | TCC | CTT | ATT | GCC |
| Val | Phe | Gly | Ser | Val | Asp | Ser | Tyr | Asp | Gly | Leu | Ala | Ser | Leu | Ile | Ala |
| AAG | GAA | TCT | GGG | ATT | GCG | GTT | ATC | TCC | GTG | GAG | TAT | AGG | CTC | GCC | CCT |
| Lys | Glu | Ser | Gly | Ile | Ala | Val | Ile | Ser | Val | Glu | Tyr | Arg | Leu | Ala | Pro |
| GAG | CAC | AAG | TTC | CCC | ACC | GCA | GTC | AAC | GAC | TCG | TGG | GAT | GCG | CTT | CTC |
| Glu | His | Lys | Phe | Pro | Thr | Ala | Val | Asn | Asp | Ser | Trp | Asp | Ala | Leu | Leu |
| TGG | ATC | GCG | GAG | AAC | GGA | GGC | AAG | CTG | GGG | CTC | GAC | ACC | TCG | AGA | CTT |
| Trp | Ile | Ala | Glu | Asn | Gly | Gly | Lys | Leu | Gly | Leu | Asp | Thr | Ser | Arg | Leu |
| GCC | GTG | GCT | GGG | GAT | AGT | GCT | GGA | GGA | AAC | CTG | TCT | GCC | GTG | GTG | TCC |
| Ala | Val | Ala | Gly | Asp | Ser | Ala | Gly | Gly | Asn | Leu | Ser | Ala | Val | Val | Ser |
| CTC | CTG | GAC | AGG | GAC | CAG | GGT | AAG | GGA | CTG | GTT | AGT | TAT | CAG | GTC | CTA |
| Leu | Leu | Asp | Arg | Asp | Gln | Gly | Lys | Gly | Leu | Val | Ser | Tyr | Gln | Val | Leu |
| ATC | TAC | CCA | GCA | GTG | AAC | ATG | GTC | GAT | AAC | TCC | CCA | TCC | GTC | AGG | GAG |
| Ile | Tyr | Pro | Ala | Val | Asn | Met | Val | Asp | Asn | Ser | Pro | Ser | Val | Arg | Glu |
| TAC | GGC | GAG | GGA | TAC | TTC | CTC | ACC | AGG | TCC | ATG | ATG | AAC | TGG | TTC | GGG |
| Tyr | Gly | Glu | Gly | Tyr | Phe | Leu | Thr | Arg | Ser | Met | Met | Asn | Trp | Phe | Gly |
| ACC | ATG | TAC | TTC | TCC | TCT | GGA | AGG | GAA | GCG | GTA | TCC | CCC | TAC | GCC | TCT |
| Thr | Met | Tyr | Phe | Ser | Ser | Gly | Arg | Glu | Ala | Val | Ser | Pro | Tyr | Ala | Ser |
| CCA | GCC | TTG | GCT | GAC | CTA | CAT | AAC | CTC | CCA | CCC | TCA | CTG | GTG | ATC | ACT |
| Pro | Ala | Leu | Ala | Asp | Leu | His | Asn | Leu | Pro | Pro | Ser | Leu | Val | Ile | Thr |
| GCA | GAG | TAT | GAT | CCC | CTA | AGG | GAT | CAG | GGA | GAG | ACC | TAC | TCT | CAC | TCC |
| Ala | Glu | Tyr | Asp | Pro | Leu | Arg | Asp | Gln | Gly | Glu | Thr | Tyr | Ser | His | Ser |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| CTA | AAC | GAG | GCT | GGA | AAC | GTA | TCA | ACC | TTG | GTT | AGA | TAT | CAA | GGA | ATG |
| Leu | Asn | Glu | Ala | Gly | Asn | Val | Ser | Thr | Leu | Val | Arg | Tyr | Gln | Gly | Met |
| ATT | CAC | GGC | TTC | CTG | TCC | TTC | TAC | GAG | TGG | ATA | ACT | GCC | GGT | AAA | CTA |
| Ile | His | Gly | Phe | Leu | Ser | Phe | Tyr | Glu | Trp | Ile | Thr | Ala | Gly | Lys | Leu |
| GCC | ATT | CAC | CAC | ATT | GCT | GGG | GTT | CTG | AGA | TCT | GTC | CTT | TA  |     |     |
| Ala | Ile | His | His | Ile | Ala | Gly | Val | Leu | Arg | Ser | Val | Leu |     |     |     |

**Figure 14**  
**Thermotoga neapolitana 5068 Esterase 56mc4**

```

GTG GCC TTC TTC GAT ATG CCC CTT GAG GAA CTG AAA AAG TAC CGG CCT
Val Ala Phe Phe Asp Met Pro Leu Glu Glu Leu Lys Lys Tyr Arg Pro
GAA AGG TAC GAG GAG AAA GAT TTC GAT GAG TTC TGG AGG GAA ACA CTT
Glu Arg Tyr Glu Glu Lys Asp Phe Asp Glu Phe Trp Arg Glu Thr Leu
AAA GAA AGC GAA GGA TTC CCT CTG GAT CCC GTC TTT GAA AAG GTG GAC
Lys Glu Ser Glu Gly Phe Pro Leu Asp Pro Val Phe Glu Lys Val Asp
TTT CAT CTC AAA ACG GTT GAA ACG TAC GAT GTT ACT TTC TCT GGA TAC
Phe His Leu Lys Thr Val Glu Thr Tyr Asp Val Thr Phe Ser Gly Tyr
AGG GGG CAG AGA ATA AAG GGC TGG CTT CTT GTT CCG AAG TTG GCG GAA
Arg Gly Gln Arg Ile Lys Gly Trp Leu Leu Val Pro Lys Leu Ala Glu
GAA AAG CTT CCA TGC GTC GTG CAG TAC ATA GGT TAC AAT GGT GGA AGG
Glu Lys Leu Pro Cys Val Val Gln Tyr Ile Gly Tyr Asn Gly Gly Arg
GGT TTT CCA CAC GAC TGG CTG TTC TGG CCG TCA ATG GGT TAC ATC TGT
Gly Phe Pro His Asp Trp Leu Phe Trp Pro Ser Met Gly Tyr Ile Cys
TTT GTC ATG GAC ACC AGG GGG CAG GGA AGC GGC TGG ATG AAG GGA GAC
Phe Val Met Asp Thr Arg Gly Gln Gly Ser Gly Trp Met Lys Gly Asp
ACA CCG GAT TAC CCT GAG GGT CCA GTC GAT CCA CAG TAC CCC GGA TTC
Thr Pro Asp Tyr Pro Glu Gly Pro Val Asp Pro Gln Tyr Pro Gly Phe
ATG ACG AGG GGC ATT CTG GAT CCG GGA ACC TAT TAC TAC AGG CGA GTC
Met Thr Arg Gly Ile Leu Asp Pro Gly Thr Tyr Tyr Tyr Arg Arg Val
TTC GTG GAT GCG GTC AGG GCG GTG GAA GCA GCC ATT TCC TTC CCG AGA
Phe Val Asp Ala Val Arg Ala Val Glu Ala Ala Ile Ser Phe Pro Arg
GTG GAT TCC AGG AAG GTG GTG GTG GCC GGA GGC AGT CAG GGT GGG GGA
Val Asp Ser Arg Lys Val Val Val Ala Gly Gly Ser Gln Gly Gly Gly
ATC CCC CTT GCG GTG AGT GCC CTG TCG AAC AGG GTG AAG GCT CTG CTC
Ile Pro Leu Ala Val Ser Ala Leu Ser Asn Arg Val Lys Ala Leu Leu
TGC GAT GTG CCG TTT CTG TGC CAC TTC AGA AGG GCC GTG CAA CTT GTC
Cys Asp Val Pro Phe Leu Cys His Phe Arg Arg Ala Val Gln Leu Val
GAC ACA CAC CCA TAC GTG GAG ATC ACC AAC TTC CTC AAA ACC CAC AGG
Asp Thr His Pro Tyr Val Glu Ile Thr Asn Phe Leu Lys Thr His Arg
GAC AAA GAG GAG ATT GTT TTC AGA ACA CTT TCC TAC TTC GAT GGT GTG
Asp Lys Glu Glu Ile Val Phe Arg Thr Leu Ser Tyr Phe Asp Gly Val

```

```

AAC TTT GCA GCA AGG GCA AAG GTG CCC GCC CTG TTT TCC GTT GGG CTC
Asn Phe Ala Ala Arg Ala Lys Val Pro Ala Leu Phe Ser Val Gly Leu

ATG GAC ACC ATC TGT CCT CCC TCG ACG GTC TTC GCC GCT TAC AAC CAC
Met Asp Thr Ile Cys Pro Pro Ser Thr Val Phe Ala Ala Tyr Asn His
TAC GCC GGT CCA AAG GAG ATC AGA ATC TAT CCG TAC AAC AAC CAC GAA
Tyr Ala Gly Pro Lys Glu Ile Arg Ile Tyr Pro Tyr Asn Asn His Glu
GGT GGA GGT TCT TTC CAG GCA ATT GAG CAG GTG AAA TTC TTG AAG AGA
Gly Gly Gly Ser Phe Gln Ala Ile Glu Gln Val Lys Phe Leu Lys Arg
CTA TTT GAG GAA GGC TAG

```

[illegible]



**Figure 15**  
**Melittangium lichenicola Esterase 77mc1**

ATG CGC ACC CTC TCC TTC GGT CCG ATG ACC ACA GGG GGA AGC ATT CAC  
Met Arg Thr Leu Ser Phe Gly Pro Met Thr Thr Gly Gly Ser Ile His  
ATG GCG ACC ATG GAC GTG ATG CGC GGG CCG GGG ATG CAG CGG CTG TCA  
Met Ala Thr Met Asp Val Met Arg Gly Pro Gly Met Gln Arg Leu Ser  
CAG GGC GCC AGG GAG GCC GCG AAC CAC CCC TGG GCG AAG CGA CTG GGC  
Gln Gly Ala Arg Glu Ala Ala Asn His Pro Trp Ala Lys Arg Leu Gly  
CGC ATG GGC TAC GCG GCC AAG GGC GCC GTG TAC GCC ATC ATC GGC GTG  
Arg Met Gly Tyr Ala Ala Lys Gly Ala Val Tyr Ala Ile Ile Gly Val  
CTC GCG CTG AAG CTC GCG GCG GGC GAG GGC GGC CGG ACC ACG GAC AGC  
Leu Ala Leu Lys Leu Ala Ala Gly Glu Gly Gly Arg Thr Thr Asp Ser  
CAC GGC GCG GTG AAC ACC GTG GCG CAC GGG CCC TTC GGC GTC GCG CTG  
His Gly Ala Val Asn Thr Val Ala His Gly Pro Phe Gly Val Ala Leu  
CTG GCG GTG CTG GTG GTG GGC CTG CTG GGC TAC GTG GTC TGG AGG TTC  
Leu Ala Val Leu Val Val Gly Leu Leu Gly Tyr Val Val Trp Arg Phe  
GCC CAG GCC TTC GTG GAC ACG GAG GAC AAG GGC TCC GAC GCG AAG GGA  
Ala Gln Ala Phe Val Asp Thr Glu Asp Lys Gly Ser Asp Ala Lys Gly  
ATC GCC ACG CGC GCC ATG TAC TTC CTC AGC GGC TGC ATC TAC GCG TCG  
Ile Ala Thr Arg Ala Met Tyr Phe Leu Ser Gly Cys Ile Tyr Ala Ser  
CTG GCC TTC TTC GCC GCG CAG TCC CTG GTG GGC GCC GCG CAC GGC CGG  
Leu Ala Phe Phe Ala Ala Gln Ser Leu Val Gly Ala Ala His Gly Arg  
AGC AAG GGG ACG CAG GGC TGG ACG GCC ACG CTG ATG GAG CAG CCC TTT  
Ser Lys Gly Thr Gln Gly Trp Thr Ala Thr Leu Met Glu Gln Pro Phe  
GGC CGC GTG CTG GTG GCG CTG GTG GGG CTG GGC ATC GTG GGC TTC GCG  
Gly Arg Val Leu Val Ala Leu Val Gly Leu Gly Ile Val Gly Phe Ala  
CTG AAG CAG TTC CAC ACC GCG TGG AAG GCG AAG TTC CGG GAG AAG CTC  
Leu Lys Gln Phe His Thr Ala Trp Lys Ala Lys Phe Arg Glu Lys Leu  
ACC CTC ACC GGA CTG GCT GCC CGG AAG CAG CAC CAC ATC GAG CGC ATG  
Thr Leu Thr Gly Leu Ala Ala Arg Lys Gln His His Ile Glu Arg Met  
TGC CAG TTC GGC ATC GCC GCG CGC GGC GTG GTG TTC GCC GTC ATC GGC  
Cys Gln Phe Gly Ile Ala Ala Arg Gly Val Val Phe Ala Val Ile Gly  
GGC TTC CTC GTC CGC TCC GCC GTG GAC GCG AAC CCC GGC GAG GCC AAG  
Gly Phe Leu Val Arg Ser Ala Val Asp Ala Asn Pro Gly Glu Ala Lys

GGC CTG GGA GAG GCC CTG GCC GTC GTC GCG AGG CAG CCG TCC GGC GAC  
Gly Leu Gly Glu Ala Leu Ala Val Val Ala Arg Gln Pro Ser Gly Asp  
  
GTG CTC CTG GGG GTG GTG GCG GCG GGC CTG GTG GCC TAC GCC GCC TAC  
Val Leu Leu Gly Val Val Ala Ala Gly Leu Val Ala Tyr Ala Ala Tyr  
CTG TTC CTC CAG GCG CGC TAC CGC GAA CTC TAG  
Leu Phe Leu Gln Ala Arg Tyr Arg Glu Leu

**Figure 16**  
**Whale Mat Sample 11.801 Esterase es2**

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ATG | AGC | AAA | TTC | GCA | ATA | CTC | TGG | GCG | TTG | ATA | ACG | GCA | TAC | CTG | CCG |
| Met | Ser | Lys | Phe | Ala | Ile | Leu | Trp | Ala | Leu | Ile | Thr | Ala | Tyr | Leu | Pro |
| GAA | CCT | GTG | ATG | AAA | CTG | GTA | TAT | TTA | GGG | CGG | CGC | GAA | ACG | CTT | GGG |
| Glu | Pro | Val | Met | Lys | Leu | Val | Tyr | Leu | Gly | Arg | Arg | Glu | Thr | Leu | Gly |
| GCA | CGG | ACG | CTT | GAC | GTT | AAA | GCC | CAA | GCT | GTC | GGG | CGG | CTG | GCC | AAT |
| Ala | Arg | Thr | Leu | Asp | Val | Lys | Ala | Gln | Ala | Val | Gly | Arg | Leu | Ala | Asn |
| GCA | ACA | AGA | CCT | GTC | GGG | GTG | ATT | CCG | ACG | GTC | GAG | GAA | AGC | CGG | AAG |
| Ala | Thr | Arg | Pro | Val | Gly | Val | Ile | Pro | Thr | Val | Glu | Glu | Ser | Arg | Lys |
| ATG | ACG | GAT | AAA | GCC | GTT | AGC | CTT | TTT | GAT | CAG | CCC | GCC | CCC | GAA | TTA |
| Met | Thr | Asp | Lys | Ala | Val | Ser | Leu | Phe | Asp | Gln | Pro | Ala | Pro | Glu | Leu |
| TTC | CGT | AAA | AAA | GAC | ATT | CAG | ATT | GAC | GGG | GCT | GAA | GGG | CCT | ATT | GAT |
| Phe | Arg | Lys | Lys | Asp | Ile | Gln | Ile | Asp | Gly | Ala | Glu | Gly | Pro | Ile | Asp |
| GCC | CGT | ATT | TAC | AGC | GGC | CCT | GCA | AAA | CAT | CGC | CCR | CGR | CCA | ATW | CTA |
| Ala | Arg | Ile | Tyr | Ser | Gly | Pro | Ala | Lys | His | Arg | Pro | Arg | Pro | Ile | Leu |
| GTG | TAT | TTT | CAC | GGC | GGT | GGC | TGG | GTT | CAG | GGC | AAT | CTG | GAC | AGC | CAT |
| Val | Tyr | Phe | His | Gly | Gly | Gly | Trp | Val | Gln | Gly | Asn | Leu | Asp | Ser | His |
| GAC | GGG | GTT | TGC | GGC | AAG | CTG | GCA | AAA | TGG | GCG | AAC | TGC | ATT | GTT | ATC |
| Asp | Gly | Val | Cys | Gly | Lys | Leu | Ala | Lys | Trp | Ala | Asn | Cys | Ile | Val | Ile |
| TCG | GTC | GAT | TAT | CGT | CTA | GCG | CCC | GAA | CAC | AAA | TTT | CCT | TGT | GCG | CCG |
| Ser | Val | Asp | Tyr | Arg | Leu | Ala | Pro | Glu | His | Lys | Phe | Pro | Cys | Ala | Pro |
| CTT | GAT | GCG | ATT | GCG | GCC | TAT | AAA | TGG | GTG | CGC | GCC | AAC | GCA | ACA | AAC |
| Leu | Asp | Ala | Ile | Ala | Ala | Tyr | Lys | Trp | Val | Arg | Ala | Asn | Ala | Thr | Asn |
| CTT | GGC | GGC | GAT | CCT | GAA | CGT | ATC | GGC | GTT | GGC | GGC | GAT | AGC | GCA | GGG |
| Leu | Gly | Gly | Asp | Pro | Glu | Arg | Ile | Gly | Val | Gly | Gly | Asp | Ser | Ala | Gly |
| GGC | AAT | CTT | GCC | GCC | GTT | GTC | TGC | CAA | CAA | ACC | GCC | ATG | AAC | GGC | GAG |
| Gly | Asn | Leu | Ala | Ala | Val | Val | Cys | Gln | Gln | Thr | Ala | Met | Asn | Gly | Glu |
| CGC | ACA | CCA | GAT | CTG | CAA | GTC | CTG | ATC | TAT | CCG | GCG | CTG | GAT | GCA | CGC |
| Arg | Thr | Pro | Asp | Leu | Gln | Val | Leu | Ile | Tyr | Pro | Ala | Leu | Asp | Ala | Arg |
| ATG | ATC | TCG | ACC | TCG | ATG | GAG | GAA | TTG | CGT | GAT | GCC | TAC | ATC | TTG | CCG |
| Met | Ile | Ser | Thr | Ser | Met | Glu | Glu | Leu | Arg | Asp | Ala | Tyr | Ile | Leu | Pro |
| AAA | TCC | AGA | ATG | GAG | TAT | TTC | CTC | GGC | CTA | TAT | ACG | CGT | GGC | CCT | GAC |
| Lys | Ser | Arg | Met | Glu | Tyr | Phe | Leu | Gly | Leu | Tyr | Thr | Arg | Gly | Pro | Asp |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| GAT | ATC | GAG | GAC | CTT | AGG | ATG | TCG | CCA | ATT | CTC | AGG | GAT | ACC | GTC | GCG |
| Asp | Ile | Glu | Asp | Leu | Arg | Met | Ser | Pro | Ile | Leu | Arg | Asp | Thr | Val | Ala |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| GAT | CAA | CCC | CAA | GCC | TGC | ATT | GTC | ACC | TGT | GGG | TTT | GAC | CCT | GCG | CGA |
| Asp | Gln | Pro | Gln | Ala | Cys | Ile | Val | Thr | Cys | Gly | Phe | Asp | Pro | Ala | Arg |
| CGA | CGG | GAA | CAC | CTA | CGC | CGA | ACG | CTT | AAT | TGC | CGA | GGG | GAT | AGA | CGT |
| Arg | Arg | Glu | His | Leu | Arg | Arg | Thr | Leu | Asn | Cys | Arg | Gly | Asp | Arg | Arg |

TA

Figure 17  
Whale Mat Sample AD3059 Esterase es4

GTG AGC ATT CGT CTG CGA CTG TTA AAC TGG TTT TTG AAT ACC TTT GAA  
Val Ser Ile Arg Leu Arg Leu Leu Asn Trp Phe Leu Asn Thr Phe Glu  
AAA CCA AAA CTG GCC GCG GCC AAA ACG CCG GAT GAT TTG CGA AAA TCG  
Lys Pro Lys Leu Ala Ala Ala Lys Thr Pro Asp Asp Leu Arg Lys Ser  
TTT GAA TTA AAG GCG AGG TTT TTG TTT CCG GCG CCA CGT AAA ACA AGG  
Phe Glu Leu Lys Ala Arg Phe Leu Phe Pro Ala Pro Arg Lys Thr Arg  
TTT AGT CAT GAT GTA TTG CAG TCA GGC ATC GGG TCG GTA AAT GCC CAG  
Phe Ser His Asp Val Leu Gln Ser Gly Ile Gly Ser Val Asn Ala Gln  
TGG GCG AAA TCC AAA TCT GCA TCT GAT GAC AGG GTA ATC CTG TAT TTT  
Trp Ala Lys Ser Lys Ser Ala Ser Asp Asp Arg Val Ile Leu Tyr Phe  
CAT GGG GGA GGG TAT GTT TTT GGG TCA CCA AAA ACG CAC CGT GCA ATG  
His Gly Gly Gly Tyr Val Phe Gly Ser Pro Lys Thr His Arg Ala Met  
TTG GCG CGC TTG TCG GCA ATG ACA GGT CTT TCT GCG TGC CTT CCA GAT  
Leu Ala Arg Leu Ser Ala Met Thr Gly Leu Ser Ala Cys Leu Pro Asp  
TAT AGG TTG GCA CCA GAG CAC CCA TTT CCA GCC GCG ATC GAA GAT GCA  
Tyr Arg Leu Ala Pro Glu His Pro Phe Pro Ala Ala Ile Glu Asp Ala  
GTT TTA TCG TAT AAA TGT TTA CTA GAG CGA GCA ATC GAG CCC CAA AAT  
Val Leu Ser Tyr Lys Cys Leu Leu Glu Arg Ala Ile Glu Pro Gln Asn  
ATT ATA CTG GGG GGG GAC AGT GCT GGT GGC GGT TTG GTT CTT GCT TTG  
Ile Ile Leu Gly Gly Asp Ser Ala Gly Gly Gly Leu Val Leu Ala Leu  
CTT GCA GAA ATC AAG GCC CAA TCC TTG CCC AAA CCT GCT GGC GTT TTT  
Leu Ala Glu Ile Lys Ala Gln Ser Leu Pro Lys Pro Ala Gly Val Phe  
GCC TTG TCG CCT TTG GTT GAT TTA TCA TTT TCG GGC CTT TCG TTT TCT  
Ala Leu Ser Pro Leu Val Asp Leu Ser Phe Ser Gly Leu Ser Phe Ser  
AAA AAT GCC CAA ACC GAT GTG ATG TTG CCC GCA TCA CGG GCT GCG GAT  
Lys Asn Ala Gln Thr Asp Val Met Leu Pro Ala Ser Arg Ala Ala Asp  
ATG GCG ACC TTG TAT TTG GAT GGG GCC GAT GCA GAT GAT CCA CGT GCA  
Met Ala Thr Leu Tyr Leu Asp Gly Ala Asp Ala Asp Asp Pro Arg Ala  
TCG CCG CTG CAG GCG GAT TTT TCT GGC ATG CCG CCT GTA TTT CTG ACA  
Ser Pro Leu Gln Ala Asp Phe Ser Gly Met Pro Pro Val Phe Leu Thr  
GCA AGT GAC AGT GAA ATC CTG TTG GAT GAT TGC CTG CGG ATG GCG GAT  
Ala Ser Asp Ser Glu Ile Leu Leu Asp Asp Cys Leu Arg Met Ala Asp

CAC TTG CGT GCG CAA GGT GTC GTT GTG ACA GAC CGG ATT GTT GAA AAC  
His Leu Arg Ala Gln Gly Val Val Val Thr Asp Arg Ile Val Glu Asn  
CAT CCA CAT GTT TGG CAT ATT TTT CAA CGC CTT CTA CCC GAA GCA GAT  
His Pro His Val Trp His Ile Phe Gln Arg Leu Leu Pro Glu Ala Asp  
CAG GGG CTG CGG GCG ATT GCC GCG TGG ATT AAA CCT CTT TTA TCA GGT  
Gln Gly Leu Arg Ala Ile Ala Ala Trp Ile Lys Pro Leu Leu Ser Gly  
TCA AAC GAA AGC TA  
Ser Asn Glu Ser

**Figure 18**  
***Microscilla furvescens* Esterase 53sc2**

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ATG | CTT | ACA | TTT | AAT | GTT | TTA | TAT | GGT | ATG | ATG | AAA | CAA | AAA | CTA | GCA |
| Met | Leu | Thr | Phe | Asn | Val | Leu | Tyr | Gly | Met | Met | Lys | Gln | Lys | Leu | Ala |
| GCA | ATT | CTC | ATG | TTT | TTA | GGG | CTA | TCA | GCA | GCA | GAG | GCT | CAA | GAC | TGG |
| Ala | Ile | Leu | Met | Phe | Leu | Gly | Leu | Ser | Ala | Ala | Glu | Ala | Gln | Asp | Trp |
| CCT | GAC | CTA | CAG | AAA | TAT | CGT | AGT | GCT | AAT | AAA | GAA | GCC | AAA | TTA | CTT |
| Pro | Asp | Leu | Gln | Lys | Tyr | Arg | Ser | Ala | Asn | Lys | Glu | Ala | Lys | Leu | Leu |
| CCA | AAG | GAA | AAC | CGG | AAG | GTG | GTT | TTT | ATG | GGC | AAC | TCC | ATT | ACA | GAA |
| Pro | Lys | Glu | Asn | Arg | Lys | Val | Val | Phe | Met | Gly | Asn | Ser | Ile | Thr | Glu |
| GCC | TGG | ATT | AGT | CAG | CGA | CCT | GAG | TTT | TTT | AGT | GAA | AAT | GGG | TTT | ATC |
| Ala | Trp | Ile | Ser | Gln | Arg | Pro | Glu | Phe | Phe | Ser | Glu | Asn | Gly | Phe | Ile |
| GGT | CGA | GGC | ATC | AGT | GGC | CAG | ACA | ACC | CCT | CAG | ATG | TTG | TTG | AGA | TTC |
| Gly | Arg | Gly | Ile | Ser | Gly | Gln | Thr | Thr | Pro | Gln | Met | Leu | Leu | Arg | Phe |
| CGA | CAG | GAT | GTG | ATA | GAC | CTG | CAG | CCA | AAG | GCT | GTA | GTG | ATA | CTA | GCT |
| Arg | Gln | Asp | Val | Ile | Asp | Leu | Gln | Pro | Lys | Ala | Val | Val | Ile | Leu | Ala |
| GGT | ACC | AAT | GAC | GTA | GCT | CAA | AAT | ACC | GGG | CCG | ATG | ACC | ATT | GAG | GAA |
| Gly | Thr | Asn | Asp | Val | Ala | Gln | Asn | Thr | Gly | Pro | Met | Thr | Ile | Glu | Glu |
| TCG | CTT | GCT | AAC | ATT | AAG | TCT | ATG | GTG | GAG | CTG | GCG | CAA | GCC | AAT | GGG |
| Ser | Leu | Ala | Asn | Ile | Lys | Ser | Met | Val | Glu | Leu | Ala | Gln | Ala | Asn | Gly |
| ATC | ACG | CCT | GTT | TTG | TGT | ACC | GTG | CTG | CCT | GCA | GAT | CGT | TTC | AGC | TGG |
| Ile | Thr | Pro | Val | Leu | Cys | Thr | Val | Leu | Pro | Ala | Asp | Arg | Phe | Ser | Trp |
| CGA | CCT | GAG | CTT | ACA | CCC | GCA | GAA | ACT | ATC | ATT | GCC | CTC | AAT | CAG | CTC |
| Arg | Pro | Glu | Leu | Thr | Pro | Ala | Glu | Thr | Ile | Ile | Ala | Leu | Asn | Gln | Leu |
| ATT | AAG | CAA | TAT | GCC | GAG | GCA | CAG | GGC | CTG | GCC | CTG | GTG | GAT | TAT | CAT |
| Ile | Lys | Gln | Tyr | Ala | Glu | Ala | Gln | Gly | Leu | Ala | Leu | Val | Asp | Tyr | His |
| GCT | GCA | CTC | ACC | AAT | AAA | GGT | GGA | GGA | CTT | CCG | GTG | AAA | TAC | GGA | GAA |
| Ala | Ala | Leu | Thr | Asn | Lys | Gly | Gly | Gly | Leu | Pro | Val | Lys | Tyr | Gly | Glu |
| GAT | GGT | GTG | CAT | CCA | AAT | GTA | GCA | GGC | TAT | CAG | GTG | ATG | GAA | AAC | ATT |
| Asp | Gly | Val | His | Pro | Asn | Val | Ala | Gly | Tyr | Gln | Val | Met | Glu | Asn | Ile |
| GTT | TTA | CCG | GTC | ATT | TCC | AGC | GAG | TTG | GCA | AAG | CTG | AAG | TA  |     |     |
| Val | Leu | Pro | Val | Ile | Ser | Ser | Glu | Leu | Ala | Lys | Leu | Lys |     |     |     |



[illegible]

Figure 20  
*Polyangium brachysporum* Esterase 78mc1

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| TTG | AAG | TAC | TTC | AAA | GCC | CGG | CTT | GCC | GGC | ATC | ACC | TTG | CTC | GGC | CTG |
| Leu | Lys | Tyr | Phe | Lys | Ala | Arg | Leu | Ala | Gly | Ile | Thr | Leu | Leu | Gly | Leu |
| CTG | GCC | TGC | ACC | TCG | GCC | TCG | GCG | CAG | ACC | GAG | CCC | ATC | GTG | TTC | GTG |
| Leu | Ala | Cys | Thr | Ser | Ala | Ser | Ala | Gln | Thr | Glu | Pro | Ile | Val | Phe | Val |
| CAC | GGC | TAT | TCC | GGC | AGC | GCA | TCC | AAC | TGG | GAC | ACC | ATG | CTG | GGC | CGC |
| His | Gly | Tyr | Ser | Gly | Ser | Ala | Ser | Asn | Trp | Asp | Thr | Met | Leu | Gly | Arg |
| TTC | CGG | TCG | AAC | GGT | TAT | GCG | TCC | GGC | TCG | CTC | TAC | ACC | TTC | AAC | TAC |
| Phe | Arg | Ser | Asn | Gly | Tyr | Ala | Ser | Gly | Ser | Leu | Tyr | Thr | Phe | Asn | Tyr |
| AAC | TCG | TTG | GTC | AGC | AGC | AAC | CGC | ACC | AGC | GCC | AGC | GAG | CTG | CGC | AGC |
| Asn | Ser | Leu | Val | Ser | Ser | Asn | Arg | Thr | Ser | Ala | Ser | Glu | Leu | Arg | Ser |
| TTC | GTC | AAC | ACC | GTG | CGT | TCG | CGC | CAC | GGC | AAC | GCC | CGC | ATC | GCG | CTG |
| Phe | Val | Asn | Thr | Val | Arg | Ser | Arg | His | Gly | Asn | Ala | Arg | Ile | Ala | Leu |
| GTC | GCC | CAC | TCC | AAC | GGC | GGG | CTG | GTG | TCG | CGC | TGG | TAT | CGC | GCG | GAG |
| Val | Ala | His | Ser | Asn | Gly | Gly | Leu | Val | Ser | Arg | Trp | Tyr | Arg | Ala | Glu |
| CTG | GGC | GGC | GAA | ACG | GCC | ACC | CGC | CGC | TTC | GTG | ACG | CTG | GGC | ACG | CCG |
| Leu | Gly | Gly | Glu | Thr | Ala | Thr | Arg | Arg | Phe | Val | Thr | Leu | Gly | Thr | Pro |
| CAC | CGG | GGC | ACC | ACC | TGG | GCC | TAT | GCG | TGC | TAC | AGC | CCC | GCA | TGT | TTC |
| His | Arg | Gly | Thr | Thr | Trp | Ala | Tyr | Ala | Cys | Tyr | Ser | Pro | Ala | Cys | Phe |
| GAG | ATG | CGC | CCC | GGC | TCC | AGC | TTG | CTG | ACC | ACG | CTG | GGC | TCG | CGT | GCC |
| Glu | Met | Arg | Pro | Gly | Ser | Ser | Leu | Leu | Thr | Thr | Leu | Gly | Ser | Arg | Ala |
| TGC | GAC | CGC | TCG | CTG | TGG | TCG | AAC | ACC | GAC | GGC | ATC | ATC | CTG | CCG | GCG |
| Cys | Asp | Arg | Ser | Leu | Trp | Ser | Asn | Thr | Asp | Gly | Ile | Ile | Leu | Pro | Ala |
| TCC | AGC | GCG | CAG | TGT | GGT | GTC | AGC | ACG | CGC | ACT | GCC | GAC | GTC | AGC | CAT |
| Ser | Ser | Ala | Gln | Cys | Gly | Val | Ser | Thr | Arg | Thr | Ala | Asp | Val | Ser | His |
| CTC | GAC | CTG | CTG | ACC | GAC | TCT | CGC | GTG | TAC | ACG | CAG | TTG | CGC | ACG | CAG |
| Leu | Asp | Leu | Leu | Thr | Asp | Ser | Arg | Val | Tyr | Thr | Gln | Leu | Arg | Thr | Gln |
| TTG | CAA | TGA | GGG | TGA | CGG | TGC | ACC | GAA | CGT | GCA | CCT | G   |     |     |     |
| Leu | Gln | End | Gly | End | Arg | Cys | Thr | Glu | Arg | Ala | Pro |     |     |     |     |

6644666 = 6446666

**DECLARATION AND POWER OF ATTORNEY**

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

**ESTERASES**

the specification of which [ ] is attached hereto or [X] was filed on February 16, 1996 as Application Serial No. 08/602,359 and was amended on \_\_\_\_\_ (if applicable).

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the patentability of this application in accordance with Title 37, Code of Federal Regulations, Section 1.56(a).

I hereby claim foreign priority benefits under Title 35, United States Code, Section 119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed. Prior Foreign Application(s):

**Priority Claimed**

Yes      No

☐      ☐

| (Number) | (Country) | (Day/Month/Year Filed) |
|----------|-----------|------------------------|
|----------|-----------|------------------------|

I hereby claim the benefit under Title 35, United States Code, Section 120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, Section 112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, Section 1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

| (Application Serial No.) | (Filing Date) | <u>Pending</u><br>(Status - patented, pending, abandoned) |
|--------------------------|---------------|---|
| _____                    | _____         | _____   |
| (Application Serial No.) | (Filing Date) | (Status - patented, pending, abandoned)                   |

I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith: John N. Bain (Reg. No. 18,651); John G. Gilfillan, III (Reg. No. 22,746); Elliot M. Olstein (Reg. No. 24,025); Raymond J. Lillie (Reg. No. 31,778); Charles J. Herron (Reg. No. 28,019); William Squire (Reg. No. 25,378); Kenneth S. Weitzman (Reg. No. 36,306); and Gregory Ferraro (Reg. No. 36,134). Address correspondence and telephone calls to Charles J. Herron c/o Carella, Byrne, Bain, Gilfillan, Cecchi, Stewart & Olstein, 6 Becker Farm Road, Roseland, NJ 07068 - (201) 994-1700.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of inventor: Dan E. Robertson

Inventor's signature: Dan E. Robertson Date: May 8, 1996

Residence: 33 Evergreen Lane, Haddonfield, NJ 08033 Citizenship: United States

Post Office Address: same

Full name of inventor: Dennis Murphy

Inventor's signature: Dennis J. Murphy Date: May 8, 1996

Residence: 10 Fairway Road, Paoli, PA 19301 Citizenship: United States

Post Office Address: same

Full name of inventor: John Reid

Inventor's signature: John C. Reid Date: May 8, 1996

Residence: 922 Montgomery Avenue Apt. J-2, Bryn Mawr, PA 19010 Citizenship: United States

Post Office Address: same



Full name of inventor: Anthony M. Maffia  
Inventor's signature: [Signature] Date: 5/5/96  
Residence: 2505-3A Cedar Tree Drive, Wilmington, DE 19810 Citizenship: United States  
Post Office Address: same

Full name of inventor: Steven Link  
Inventor's signature: [Signature] Date: 5/8/96  
Residence: 108 Gibson Avenue, Wilmington, DE 19803 Citizenship: United States  
Post Office Address: same

Full name of inventor: Ronald V. Swanson  
Inventor's signature: [Signature] Date: 5/8/96  
Residence: 309 N. Lemon Street, Apt. A, Media, PA 19063 Citizenship: United States  
Post Office Address: same

Full name of inventor: Patrick V. Warren  
Inventor's signature: [Signature] Date: 5/8/96  
Residence: 35507 Sheffield Ave., Philadelphia, PA 19163 Citizenship: United States  
Post Office Address: same

Full name of inventor: Anna Kosmotka  
Inventor's signature: [Signature] Date: 5/8/96  
Residence: 224 Harrison Road, Brookhaven, PA 19015 Citizenship: United States  
Post Office Address: same




In re Application of:  
Robertson et al.  
Application No.: Unassigned  
Filed: Herewith  
Page 2

as associate attorney of record to prosecute this application as well as any continuation and divisional applications and to transact all business in the Patent and Trademark Office in connection therewith.

Respectfully submitted,

Date: August 24, 1999



Lisa A. Haile, Ph.D.

Telephone: (858) 677-1456

Facsimile: (858) 677-1465

GRAY, CARY, WARE & FREIDENRICH, LLP  
4365 Executive Drive, Suite 1600  
San Diego, California 92121-2189